

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 14:23:59 ; Search time 8119.47 Seconds
(without alignments)
11509.077 Million cell updates/sec

Title: US-09-892-635A-44
Perfect score: 2156
Sequence: 1 ggaatcccaacttttaggaat.....agcctctcattcactatgg 2156

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2125.6	98.6	4924	6	BD080859	BD080859 Banana pr
c 2	185.2	8.6	298	8	AF399940	AF399940 Musa acum
c 3	127.6	5.9	282699	8	AF506028	AF506028 Poncirus
c 4	90.2	4.2	99606	8	AC116033	AC116033 Zea mays
c 5	87.8	4.1	86019	2	AC133571	AC133571 Medicago
c 6	87.8	4.1	121952	8	AC146745	AC146745 Medicago
c 7	85.6	4.0	132605	8	AP004957	AP004957 Medicago
c 8	83.8	3.9	109284	2	AC146910	AC146910 Medicago
c 9	83.6	3.9	125883	8	AC124959	AC124959 Medicago
c 10	83.2	3.9	141590	2	AC126788	AC126788 Medicago
c 11	83	3.8	141590	2	AC126788	AC126788 Medicago
c 12	81.2	3.8	100810	8	AF411804	AF411804 Lycopersi
c 13	81	3.8	122497	8	AC087797	AC087797 Oryza sat
c 14	81	3.8	157436	2	AC144740	AC144740 Oryza sat
c 15	80.6	3.7	135301	8	AP003198	AP003198 Oryza sat
c 16	80.6	3.7	142320	8	AP003199	AP003199 Oryza sat
c 17	80.4	3.7	167484	2	AC145225	AC145225 Zea mays
c 18	80.2	3.7	144000	8	AP005247	AP005247 Oryza sat
c 19	79.8	3.7	88947	2	AP005463	AP005463 Oryza sat
c 20	79.8	3.7	106632	8	AP004500	AP004500 Lotus cor
c 21	78.8	3.7	129921	8	AC146568	AC146568 Medicago
c 22	78.4	3.6	135983	2	AC144538	AC144538 Medicago
c 23	78.4	3.6	136470	2	AC145222	AC145222 Medicago
c 24	78	3.6	177140	2	AP005612	AP005612 Oryza sat
c 25	77.8	3.6	87613	8	AP004857	AP004857 Oryza sat
c 26	77.8	3.6	107580	8	AP005590	AP005590 Oryza sat
c 27	77.6	3.6	133759	8	OSUN00289	AL731642 Oryza sat
c 28	77.4	3.6	118733	2	AP003885	AP003885 Oryza sat
c 29	77.4	3.6	120419	8	AP005734	AP005734 Oryza sat
c 30	77.4	3.6	143515	8	AP003252	AP003252 Oryza sat
c 31	77.4	3.6	152973	8	AP004463	AP004463 Oryza sat
c 32	77.2	3.6	146408	8	AC130610	AC130610 Oryza sat
c 33	76.6	3.6	97579	2	AC146722	AC146722 Medicago
c 34	76.4	3.5	74462	8	AC130731	AC130731 Oryza sat
c 35	76.4	3.5	140553	8	AC091666	AC091666 Oryza sat
c 36	76.4	3.5	161250	8	AC079634	AC079634 Genomic S
c 37	76.4	3.5	300029	8	AE017080	AE017080 Oryza sat
c 38	76.2	3.5	58293	8	AC096781	AC096781 Oryza sat
c 39	76.2	3.5	162475	2	AC145272	AC145272 Oryza sat
c 40	76.2	3.5	168267	8	AP003535	AP003535 Oryza sat
c 41	76.2	3.5	199662	2	AC146338	AC146338 Oryza sat
c 42	76.2	3.5	306161	8	AE017102	AE017102 Oryza sat
c 43	76	3.5	110000	8	AC145127_12	Continuation (13 O
c 44	76	3.5	137015	8	AC113339	AC113339 Genomic s
c 45	76	3.5	144778	8	OSUN00146	AL662935 Oryza sat

ALIGNMENTS

RESULT 1
BD080859
LOCUS
DEFINITION BD080859 4924 bp DNA linear PAT 27-AUG-2002
Banana proteins, DNA, and DNA regulatory elements associated with
fruit development.
ACCESSION BD080859
VERSION BD080859.1 GI:22626462
KEYWORDS JP 2001517446-A/9.
SOURCE Musa acuminata
ORGANISM Musa acuminata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Zingiberales; Musaceae;
Musa.
REFERENCE 1 (bases 1 to 4924)
AUTHORS May, G. and Clendennen, S.

TITLE	Banana proteins, DNA, and DNA regulatory elements associated with fruit development
JOURNAL	Patent: JP 2001517446-A 9 09-OCT-2001;
COMMENT	BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH INC OS Musa acuminata (banana)

PN	JP 2001517446-A/9
PD	09-OCT-2001
PF	23-SEP-1998 JP 2000513960
PR	25-SEP-1997 US 60/060062
PI	GREGORY MAY STEPHANIE CLENDENEN
PC	C12N15/09, A01H5/00, C07K14/415, C07K14/42, C07K14/825, C12N5/10, C12N9/10, C12N9/00, C12N9/24, C12N15/00, C12N5/00
PC	C12N9/10, C12N9/24, C12N15/00, C12N5/00
CC	Banana proteins, DNA, and DNA regulatory elements associated with fruit
CC	

```

CC      development
FH      Key
FT      Location/Qualifiers
FT      1..4924
FT      /organism='Musa acuminata (banana)'.
FEATURES
        source
        1..4924
        /organism="Musa acuminata"
        /mol_type="genomic DNA"
        /db_xref="taxon:4641"

```

ORIGIN

Query Match	98.6%;	Score 2125.6;	DB 6;	Length 4924;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 2152;	Conservative 0;	Mismatches 0;	Indels 5;	Gaps 2

Qy	1	GGATCCCAACTTTTAGGAATGGATCTTAAATATTTTAGTTATAAGTTCAAAGTTAGAAAA	60
Db	1	GGATCCCAACTTTTAGGAATGGATCTTAAATATTTTAGTTATAAGTTCAAAGTTAGAAAA	60
Qy	61	TCCTTACCAAGAGCTTTGAGPCCATGATGACATCCGTTGAAACGGGTGACATGTCTCCGA	120
Db	61	TCCTTACCAAGAGCTTTGAGTCCATGATGACATCCGTTGAAACGGGTGACATGTCTCCGA	120
Qy	121	TGACTCACATTTGGTTTCATTTCGGAAGAGTTCGGAAGAGTGCATTAAGATATTGATTTTGG	180
Db	121	TGACTCACATTTGGTTTCATTTCGGAAGAGTTCGGAAGAGTGCATTAAGATATTGATTTTGG	180
Qy	181	ATTCTTTCACTCGGTGCTTCATGAGTGACCTCAAGAGTCTCTCAAAATATCAAAAG	240
Db	181	ATTCTTTCACTCGGTGCTTCATGAGTGACCTCAAGAGTCTCTCAAAATATCAAAAG	240
Qy	241	CCGAATCACAAATGAAATGTGATTGAATTCATTTTGTCTAATGCACAAAACAGGGCAT	300
Db	241	CCGAATCACAAATGAAATGTGATTGAATTCATTTTGTCTAATGCACAAAACAGGGCAT	300
Qy	301	TCATAGCCTTTTGTTTTAAAGCAAAAACATTCCTCTCCGATTCATCCCATTCGGTCTATCG	360
Db	301	TCATAGCCTTTTGTTTTAAAGCAAAAACATTCCTCTCCGATTCATCCCATTCGGTCTATCG	360
Qy	361	GAGAGAAATTTTGAATCCATTTTCGCAATAGACCAAGCTCGAAATCCATTCGAAA	420
Db	361	GAGAGAAATTTTGAATCCATTTTCGCAATAGACCAAGCTCGAAATCCATTCGAAA	420
Qy	421	TGAGGAGATCCTCATATGAGTTTTCCAAATACATGTAATTCGACTCATTTAAACATAGTG	480
Db	421	TGAGGAGATCCTCATATGAGTTTTCCAAATACATGTAATTCGACTCATTTAAACATAGTG	480
Qy	481	GATGTGTAATGAAATGACCTCATGCSCTATCTCTCTGGGTATTAAACCAAAATATGAGA	540
Db	481	GATGTGTAATGAAATGACCTCATGCSCTATCTCTCTGGGTATTAAACCAAAATATGAGA	540
Qy	541	GTGAGCCTTCTCTGATACCAATTTGTAGGATCAGGTGGCACATAAGAGAGGGGGGAGT	600
Db	541	GTGAGCCTTCTCTGATACCAATTTGTAGGATCAGGTGGCACATAAGAGAGGGGGGAGT	600
Qy	601	GAATTAGTCAGTGGATTTAAACCTTATAAGTTTAAAAATGAATTCGTAAATACGAGAAGA	660
Db	601	GAATTAGTCAGTGGATTTAAACCTTATAAGTTTAAAAATGAATTCGTAAATACGAGAAGA	660

```

QY 1741 AACAAACATCGTACTCAATAACCCATCAGCTATAGTGTAGTACTATCTACTGTGATC 1800
Db |||||
QY 1741 AACAAACATCGTACTCAATAACCCATCAGCTATAGTGTAGTACTATCTACTGTGATC 1800
Db |||||
QY 1801 CGTACGTGAAGTGTAGCGAGTCATGATCAGGTCTGTGTCACTTATTGGCCGAACACGCTATC 1860
Db |||||
QY 1801 CGTACGTGAAGTGTAGCGAGTCATGATCAGGTCTGTGTCACTTATTGGCCGAACACGCTATC 1860
Db |||||
QY 1861 CCTATCAAAATCAGTCTTCTCAACTCTTCTAGACCTACCGCTCTCTTTTATTACTT 1920
Db |||||
QY 1861 CCTATCAAAATCAGTCTTCTCAACTCTTCTAGACCTACCGCTCTCTTTTATTACTT 1920
Db |||||
QY 1921 TTGAAGAATTCAAATCAAAACAGATACAAATAACACGGTGAGACACTGTGACATGCTA 1980
Db |||||
QY 1921 TTGAAGAATTCAAATCAAAACAGATACAAATAACACGGTGAGACACTGTGACATGCTA 1980
Db |||||
QY 1981 GTCTCTGGAAGCAATTAATTCGGCATCCACAGACGTCGTCACTTCAATCACCACATTTT 2040
Db |||||
QY 2041 TCCTACAT-ACCATGTCGCATGGCTTTGTGATGACAGACACCAAGCTTGCCTTTGG 2099
Db |||||
QY 2041 TCCTACATACCATGTCGCATGGCTTTGTGATGACAGACACCAAGCTTGCCTTTGG 2100
Db |||||
QY 2100 TTGTGCTTAACAGAGAGAGAGAGACAGACACGATAGCTCTCTCAATTCACATATGG 2156
Db |||||
QY 2101 TTGTGCTTAAC-AGAGAGAGAGAGAGACAGACGATAGCTCTCTCAATTCACATATGG 2153
Db |||||

RESULT 2
AF399940/c 298 bp DNA linear PLN 09-OCT-2003
LOCUS AF399940
DEFINITION Musa acuminata clone Radkal2 repeat sequence.
ACCESSION AF399940
VERSION AF399940.1 GI:15529477
KEYWORDS
SOURCE Musa acuminata
ORGANISM Musa acuminata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Zingiberales; Musaceae;
Musa.
REFERENCE 1 (bases 1 to 298)
AUTHORS Valarik,M., Simkova,H., Hribova,E., Safar,J., Dolezelova,M. and
Dolezel,J.
TITLE Isolation, characterization and chromosome localization of
repetitive DNA sequences in bananas (Musa spp.)
JOURNAL Chromosome Res. 10 (2), 89-100 (2002)
MEDLINE 21989232
PUBMED 11993938
REFERENCE 2 (bases 1 to 298)
AUTHORS Valarik,M., Simkova,H., Hribova,E., Safar,J., Dolezelova,M. and
Dolezel,J.
TITLE Direct Submission
SUBMITTED (16-JUL-2001) Laboratory of Molecular Cytogenetics and
Cytometry, Institute of Experimental Botany, Sokolovska 6, Olomouc
772 00, Czech Republic
JOURNAL
FEATURES
source Location/Qualifiers
1. .298
/organism="Musa acuminata"
/mol_type="genomic DNA"
/cultivar="Pisang Mas"
/db_xref="taxon:4641"
/clone="Radkal2"

repeat_region 1. .298

ORIGIN
Query Match 8.6%; Score 185.2; DB 8; Length 298;
Best Local Similarity 89.8%; Pred. No. 7e-36;
Matches 221; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

QY 6 CCAACTTTTGAAGGATGATCTTAAATTTTATAGTTTAAAGTTCAAAGTTAGAAAATCTTT 65
Db |||||
QY 246 CCAACTTTTGAAGGATGATCTTAAATTTTATAGTTTAAAGTTCAAAGTTAGAAAATCTTT 188
Db |||||

```

```

QY 66 ACCAAGAGCTTTGAGTCCATTGATGATCCGTGAAC-CGTGTACATGTCTCGATGGA 124
Db |||||
QY 187 ACCAAGAGCTTTGAGTCCATTGATGATCCGTGAACCGGTGTACATGTCTCGATGGA 128
Db |||||
QY 125 CTCACTTGGTTTCATTTCGGAAGAGTTTCAAGAGTGCATGAAGATATTGATTTTGGATTC 184
Db |||||
QY 127 CTCACTTGGTTTCATTTCGGAAGAGTTTCAAGAGTGCATGAAGATATTGATTTTGGATTC 68
Db |||||
QY 185 TTTCACTCGTTCGTCCTTTCATGAGTACCTCAAGAGTCTCTCAATATCAAAAGCCGA 244
Db |||||
QY 67 TTTCACTCGTTCGTCCTTTCATGAGTACCTCTAGAGTTCTCCAAATATCAAAAGCCGA 8
Db |||||
QY 245 ATCACA 250
Db |||||
QY 7 GTCACA 2
Db |||||

RESULT 3
AF506028/c 282699 bp DNA linear PLN 21-FEB-2003
LOCUS AF506028
DEFINITION Poncirus trifoliata citrus tristeza virus resistance gene locus,
complete sequence.
ACCESSION AF506028
VERSION AF506028.1 GI:24461847
KEYWORDS
SOURCE Poncirus trifoliata
ORGANISM Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
REFERENCE 1 (sites)
AUTHORS Yang,Z.-N., Ye,X.-R., Choi,S., Molina,J., Moonan,F., Wing,R.A.,
Roose,M.B. and Mirkov,T.E.
TITLE Construction of a 1.2-Mb contig including the citrus tristeza virus
resistance gene locus using a bacterial artificial chromosome
library of Poncirus trifoliata (L.) Raf
Genome 44 (3), 382-393 (2001)
JOURNAL 21337599
MEDLINE 11444697
PUBMED
REFERENCE 2 (bases 1 to 282699)
AUTHORS Yang,Z.-N., Ye,X.-R., Molina,J., Roose,M.L. and Mirkov,T.E.
TITLE Sequence Analysis of a 282-kilobase Region Surrounding the Citrus
Tristeza Virus Resistance Gene (Ctv) Locus in Poncirus trifoliata
L. Raf
JOURNAL Plant Physiol. 131 (2), 482-492 (2003)
MEDLINE 22474771
PUBMED 12586873
REFERENCE 3 (bases 1 to 282699)
AUTHORS Yang,Z.-N., Ye,X.-R., Molina,J., Roose,M.L. and Mirkov,T.E.
TITLE Direct Submission
SUBMITTED (24-APR-2002) Dept. of Plant Pathology and Microbiology,
The Texas A&M University Agricultural Experiment Station, 2415 E.
Hwy 83, Weslaco, TX 78596, USA
JOURNAL
FEATURES
source Location/Qualifiers
1. .282699
/organism="Poncirus trifoliata"
/mol_type="genomic DNA"
/cultivar="Pomeroy"
/db_xref="taxon:37690"
/notes="derived from BAC clones 27A14, 20J24, 83D17, and
84F5"
<306. .>5092
/genes="CTV.1"
join(<306. .377,493. .534,931. .972,2377. .2499,2870. .2926,
3477. .3614,4240. .4281,4312. .4379,5057. .>5092)
/genes="CTV.1"
/product="CTV.1"
join(<306. .377,493. .534,931. .972,2377. .2499,2870. .2926,
3477. .3614,4240. .4281,4312. .4379,5057. .5092)
/genes="CTV.1"
/notes="similar to F-box protein"
/codon_start=3

```

```
/product="CTV.1"
/protein_id="AAN62335.1"
/db_xref="GI:24461848"
/translation="IGSLFYLNLRQCSDGCKFNKLGNEITDECLVVHLSCGI
DEGLVNYLFTGISDGLRKLGLSLKSLNLDARQITDTGLAALTAHJDLFGARI
TDSGAAYLRNKLRLIEICGGGLTDAGVKHKKDLSLALNLSQNCNLTKTLEYVS
NSRITAGLRPLESCVKYANDIKRLQSRDLPNLFLTHPTTSG"
repeat_region
4157..4166
/rpt_unit="at"
repeat_region
6640..6651
/rpt_unit="ta"
gene
complement(<7554..>15436)
/gene="CTV.2"
CDS
complement(join(<7554..7764,8721..8887,8973..9161,
9491..9586,9671..9799,9881..10060,10225..10335,
10762..10908,11100..11195,11277..11435,11518..11752,
11882..12009,12143..12238,12326..12421,12525..12598,
12686..12755,12863..12972,13107..13272,13405..13514,
13627..13753,13839..14016,14547..14681,14780..14893,
14993..15172,15357..>15436))
/gene="CTV.2"
/product="CTV.2"
complement(join(7554..7764,8721..8887,8973..9161,
9491..9586,9671..9799,9881..10060,10225..10335,
10762..10908,11100..11195,11277..11435,11518..11752,
11882..12009,12143..12238,12326..12421,12525..12598,
12686..12755,12863..12972,13107..13272,13405..13514,
13627..13753,13839..14016,14547..14681,14780..14893,
14993..15172,15357..15436))
/gene="CTV.2"
/notes="similar to Arabidopsis thaliana unknown protein
with WD40 repeats"
/codon_start=1
/product="CTV.2"
/protein_id="AAN62336.1"
/db_xref="GI:24461849"
/translation="WSSLRELVLIIQLFDEEKFKETVHLEQSGFFFMKYPFDE
VFNGWDDVELKGLGFTKVDNRYSMKIFPEIRKQYLEALDKHRAKAVELVKDLK
VFNTEBELKEITQLLTLENFRENELQSKYDTPKSARSIMLVRLKLIENPLPRDK
LOFNLRKSNLWQHQCCKNRPNDIKTLFVHTCGQNGARAPSPANNELLSLPK
AGVPPGAGPFPPTAPVTPPLAGMNSPPTVTHPAVGGAGLGSPSIPAAALKH
PRTPTNFSVDPSGDHLSKTRPIGIDSEINLPVNVFSTGHSSQAFSAPED
LPKVTITLQGGSPMSMDHFVQTLILLVNTVGDIGLMEVGRRLVLRNFVWDL
GACSMPLQALVADPGVSNRVIWSPDGLFGVASHVQIYSHGGDEVROLEID
AHVGVNDIAFSPHNKOLCVITCGDDKTIKVDATNGAKQIFSGHEAPVYSCPHKH
ENIQIFSTALDGKTKAWLYNLSRVVDYERGRCTMNASDGRILFSCGTSKDG
SFIVWNESEADVKRTYQGFRRKRLVGVQFDTNKRFLAAGDPSIKPMDMSYQLLT
SIDAGGLPASPRIRFNKDGCLAVTNDNGIKILATSDGIRLLRTFENLAYDSRTS
ENSKPTTSPISAAAAAATAGLADRAASWYSIPMGNDVRSADVKPRIETESNDKS
KWKLTSLSPNOCRLRLPNLRATKISRLIFTNNGNAILALASNAIHLLWKQORTE
RNSSGKATASVAPOLWOPPGSIMMTNDVTSNPREAVPCFALSKNDSYVMSASGKIS
LFNMTEKTATFNPPPAATFLAFHPQDNNIIAIGDDSSIIQIYNNRVDEVKSLKG
HSKRITGLAFSHALNVLVSSGADSQLCVWSDGWEKQKNRFLQIPTRGTPAQSDTRV
QHQQDTHFLVHTQTAIFETTKLECVKQWVPRESSAPITHTFSCDSQLVAYACFLD
ATCVFSAALKLRCLNPSAYLPAGVSSNVHPLVIAAHQPEPNEFALGLSDGGVHP
EPLESECKWGVPPVDNGSTSSMEATPVGSGSDQQR"
repeat_region
15984..15993
/rpt_unit="tc"
repeat_region
19033..19050
/rpt_unit="aat"
repeat_region
23206..23220
/rpt_unit="aat"
repeat_region
23248..23262
/rpt_unit="aat"
repeat_region
26450..26459
/rpt_unit="at"
repeat_region
36059..36074
/rpt_unit="at"
repeat_region
37008..37031
/rpt_unit="aat"
repeat_region
37044..37053
/rpt_unit="ag"
repeat_region
37140..37155
```

```
/rpt_unit="tc"
38595..38604
/rpt_unit="at"
<39061..>40533
/gene="CTV.3"
join(<39061..40137,40222..>40533)
/gene="CTV.3"
/product="CTV.3"
join(39061..40137,40222..40533)
/gene="CTV.3"
/notes="similar to transmembrane amino acid transporter
protein"
/codon_start=1
/product="CTV.3"
/protein_id="AAN62337.1"
/db_xref="GI:24461850"
/translation="MDSNYTVIPKNSFLEQVYNDQNPSPRINSHVKIOTFDBEHDA
DSKNFVDHNNQDDHDDXPLIASKNEGSGVAGVFNLTITVIGAGIMALPAAMKV
LGLAVGVFLIIFMGIILSVSEVLLAVRSLCKATSYGEVQYALGRPAKILSEICIL
NNAGVAVVLIIVGDVMSGSLHHTGVFDQWLKGMDPDRKLLIILVAVFLPACALD
RIELSMSAASVALAVFVVVCFPIAFIKLVEKGLDPPRMSPFSGKAILDLVVI
PIMTNAVCHFNQVPIYNELEGRTPQNMHVGRUTLILVIVSSSTASIGVLLFGKDT
ESDVLTFDKDLGIGYSTALNYVRGAPLLESKRSLKALTGVLLITVYFGSTIPSI
WTAPKFGATTAVSLGFIFFPLVALRLKEGPGSLWEKFLSGLMLVLAIVVSFVGM
GNIVSLESKDD"
<42823..>45495
/gene="CTV.4"
/notes="disease resistance gene R1"
<42823..>45495
/gene="CTV.4"
/product="NBS-LRR type disease resistance protein"
42823..45495
/gene="CTV.4"
/codon_start=1
/product="NBS-LRR type disease resistance protein"
/protein_id="AAN62348.1"
/db_xref="GI:24461861"
/translation="MGNICOLISDCGAFNNRCLDFLGKAAVYSNIQDNILVALDTCL
RKLIAAKDNMRVNDARQMRRLDQVQVWVSVEVTEADAPIGDGTQBEIKLCL
GGYCKCKSGKSPFGKQVARKLRDITKIMGEGVFVADVADKPEFAPDEPTPEVVLG
OSQLEEVCLVEEPVIGLYGNGVQKTTLLTHINNKFLGSPFNLDLIVLVVSKD
LRLESIQEVTEKIGLLNDAMKSRRIEQKALDIFRIGKNFVLLDDIQWRVLDKY
GIPLNQSOTASKVVFTRSEEVCLGMAHKFKVECLSGNDAMELFRQKVGEEITLNC
HHDILEAOVTYKCGGLPLALITIGRAMACKTPEEWSVAIOVLRYSGSOFFGLNE
VYPLKXSYNLDNDRIRSCLLYCCLYPCGCCISKENLVDCWIGEGLLNGSVTJGSHE
QGVHVGILVHSCLEEDVEVMDVIRMDLWLAADCAEKENILVYAGAGUREA
PDVIEWKLRLSMENQIENLSEVPTCPHLLTLFNSDDILFNSDILFQSGMLRLKY
LNLRYMGLLVLPLIGISKLVSEYLDLSLISIPBELKALVNLKCNLYTGLRLK
IPLQILNFGRLVLRMFGNAYFSYGNVPTESVLFGGELLVEBLLGKHLEVLSLTL
GSRALQSFLTSHMLRSCRAMLLODFQGSTVDVSGLADLRLKRLRISDCELVEL
KIDYAGVQVQYGHSLQSFBVNVCSKLKDLTLLVLINLKSIEVTDCEANBELISVE
FAGNAPAFKLQYIGINLNLKSIYWKPLFPFPCLEETVSDCYELKLPUDSNAKE
HKVIRGAANWRNLQWEDEATQNAFLSCFQSLA"
complement(<46474..>51670)
/gene="CTV.5"
join(<46474..46761,48604..49006,49060..49234,49534..49620,
50490..50665,51207..>51670)
/product="monosaccharide transport protein"
join(46474..46761,48604..49006,49060..49234,49534..49620,
50490..50665,51207..51670)
/codon_start=1
/product="monosaccharide transport protein"
/protein_id="AAN62338.1"
/db_xref="GI:24461851"
/translation="WIESGAYNNYTLVQDPELLSRKRSALATGLLLVLIVPGSTMIPN
IWTAPKFGATTAVSLGFIFFPLVALRLKEGPGSLGKFLSGLMLVLAIGLLTH
SLIAKAVAAACIGAKFGIDGNPGELPKWYAIYVVLFFYIYVAGFAWGGPGLWLPVSE
ILPLKIRSAQSVNVSNMCLTFLVAQVFNKLCHLKEGLFPFGFVWMSFTYFPP
LAETKGIPIEMDMLTNDCTDNDENTTLHLKLOHLKLLRLRSPFLQDLRSQKQITG
AGGFGKNPMLNGLGRGTIALLRDRAFFLNDKLLVGMRLCNVPSGRKNTLRYSPQ
EAIFPFGCELNRRGGQSGENSSDGTETPGERKGNETIQSGJLQDHLHTRFV
TPRRLFFFTLHVDSGRHLFVYLTSMFSDSWGLSDDBELVKVLEAFNAFNCDDL
SPTSLVBSRLVDVAGMDNKIPSAMSRKRAQVSGESSIAPFSQKSNSTGFSKTSAPAL
```


TITLE		Direct Submission	
JOURNAL	Submitted (21-NOV-2002)	University of Wisconsin, Department of Horticulture, Madison, WI 53706, USA	
REFERENCE	4 (bases 1 to 99606)	Jiang,J.	
AUTHORS			
TITLE	Direct Submission		
JOURNAL	Submitted (28-NOV-2002)	University of Wisconsin, Department of Horticulture, Madison, WI 53706, USA	
REFERENCE	5 (bases 1 to 99606)	Jiang,J.	
AUTHORS			
TITLE	Direct Submission		
JOURNAL	Submitted (03-DEC-2002)	University of Wisconsin, Department of Horticulture, Madison, WI 53706, USA	
REFERENCE	6 (bases 1 to 99606)	Jiang,J.	
AUTHORS			
TITLE	Direct Submission		
JOURNAL	Submitted (04-OCT-2003)	University of Wisconsin, Department of Horticulture, Madison, WI 53706, USA	
COMMENT	On Oct 4, 2003 this sequence version replaced gi:25141017.		
FEATURES	Location/Qualifiers		
source	1..99606		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:4577"		
	/chromosome="unknown"		
	/clone="ZM15C05"		
	/note="The BAC contains DNA sequences associated with maize centromeres."		
ORIGIN			
Query Match	4.2%;	Score 90.2;	DB 8; Length 99606;
Best Local Similarity	51.9%;	Pred. NO. 1.1e-11;	
Matches 254;	Conservative	0; Mismatches 228;	Indels 7; Gaps 2;
QY	4	TCCCAACTTTAGGAATGGATCTTAAATTTTAGTTATATAGTTCAAAGTTAGAAAAATCT	63
Db	33123	TCATATTTCTTCGGAAGTGATCTAAGAAACTTGTGTGAAATCCACATCCGATACATTA	33064
QY	64	TTACCAAGAGCTTTGAGTCCATTCATGACATCCCGTGAACCGGT-GTACATGCTCCCGATG	122
Db	33063	AAACCAAGTCCTTGAGTTCATTTACTATCTCTCAACCGATTGAACATATCGGATACA	33004
QY	123	GACTCACATGGTTCCTTCATTCGGAAAAAGTTTCGAAAGAGTGCAATGAAGATATTCATTTGGAT	182
Db	33003	CTCTCATCTCTTCATATAATAAATTCGTCAAACTTCCTTTGCACACATATAGTTTGCA	32944
QY	183	TCTTTCATCGGTTGGTGGCCCTTCATGAGTAGTACCTCAAGAGTCTCCAAAATATCAAAAGCC	242
Db	32943	CTCTTCACATTTGAGTTTCCTTCATGAATTTCCATTTAGTTTATCCAAATCTCATGAGCG	32884
QY	243	GAATCACAAATGGAATGCTGATTCGAATTCATTTTGTCTATATGACAAAAACAGGCATTC	302
Db	32883	GTTGTGAGATTCTTGATACGATTGAACCCATTATATCAAGAGCATCATATAAAGACATTC	32824
QY	303	ATAGCCTTTGTGTTTAAAGCAAAAACATTCCTCTCCGATTCATCCCATTCCTCATCGGA	362
Db	32823	TGGCTTGATCATTTGTGAGGATATTCATATATCACTAACAGATGGTTCACTCTCTTC	32764
QY	363	AGAGAAAATTTTGAATTCATTTTCGACATAGACCAAGCTCGAATCATCGAAATG	422
Db	32763	AACAAA-----CAAAATCCATCCCTGACAATTTGGCCAAAATTTTCCACCATTTGCTTA	32710
QY	423	AGGAAGATCCCTCATATGAGTTTCCCAATACATGTAATTCGACTCATTAACATAGGTGGA	482
Db	32709	AGATGCATTGACATTCCTTATTTTCCAAATAATCATTAATTTGTTTCCATCAAAATGCGGTGC	32650
QY	483	TGTGTAATG 491	
Db	32649	TTTCCAATG 32641	
RESULT 5			
AC133571/c			

```

LOCUS      AC133571      86019 bp      DNA      linear      HTG 05-NOV-2003
DEFINITION Medicago truncatula clone mth2-10h12, WORKING DRAFT SEQUENCE, 2
ordered pieces.
ACCESSION  AC133571
VERSION    AC133571.28 GI:38175857
KEYWORDS   HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE     Medicago truncatula (barrel medic)
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE  1 (bases 1 to 86019)
AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE     Medicago truncatula BAC Clone mth2-10h12
JOURNAL   Unpublished
AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE     Direct Submission
JOURNAL   Submitted (14-SEP-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE  3 (bases 1 to 86019)
AUTHORS   Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE     Direct Submission
JOURNAL   Submitted (05-NOV-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT   On Nov 5, 2003 this sequence version replaced gi:380441179.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. the order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 23532: contig of 23532 bp in length
* 23533 23632: gap of unknown length
* 23633 86019: contig of 62387 bp in length.
* Location/Qualifiers
* 1..86019
* /organism="Medicago truncatula"
* /mol_type="genomic DNA"
* /db_xref="taxon:3880"
* /clone="mth2-10h12"
* /clone_lib="Medicago truncatula BAC library H2"

ORIGIN
Query Match 4.1%; Score 87.8; DB 2; Length 86019;
Best Local Similarity 51.5%; Pred. No. 4.5e-11;
Matches 226; Conservative 0; Mismatches 212; Indels 1; Gaps 1;

Qy 1 GGATCCCAACTTTTAGGAATGGATCTTAAATTTTGTATATAAGTTCAAAGTTAGAAAAA 60
Db 69956 GGTCTCCATTGAGTTGGAGACATCTTAGTATTTTCTATCTCTGCTGGTGAATTA 56897

Qy 61 TCTTTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAAACG-GTGTACATGTCCTG 119
Db 56896 GTTTTCCAGTGATCTTAGTTCAATTACTATAGATGTGAATCTAGAGTACATTTTCATCA 56837

Qy 120 ATGGACTCAGTTGGTTTCATTCGGAAAGTTCCGAAGAGTGCATAGATATGATTTTG 179
Db 56836 ATGTTTTCATCTTCAATCATTTTCAAAGATTTCAAACCTTTCTCACCAATGTCATTTCT 56777

```

```

Qy 180 GATTCTTTCTACTCGTTGGTGCCTTTCATGAGTGACCTCAAGAGTCCCTCCRAATATCAAAA 239
Db 56776 GTTTCTTTAAACATGCGCTTGTTCCCTTCATGATGGAGACTTTGAGAGATATCCACACCTTTT 56717

Qy 240 GCCGAATCACAAATTTGAAATGTGATTGAATTCATTTTGTCTAATGCACAAAAACAGGSCA 299
Db 56716 GCATATCACATTCATCACTCTTTCTACTTTTCCATGCTTAGTGACATGATAGAAG 56657

Qy 300 TTCATAGCCTTTGTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATCGCTCATC 359
Db 56556 AGTCGAGCTTTAGAAATTTAAAGATATTTTCTTTATCTCTTTTGTCCATGATGCTTCT 56597

Qy 360 GGAAGAGAAATTTTGAATCCATTTTCGACATAGACCAAGAGCTCGAAATCCATGAA 419
Db 56596 GGTGTAATTAATCTTTGTTTGTAGTAATATCATCAGTTGTTCATGTGTACATGATCCCATCT 56537

Qy 420 ATGAGGAAGATCCCTCATAT 438
Db 56536 GTTATTATTATTCACATAT 56518

RESULT 6
LOCUS    AC146745      121952 bp      DNA      linear      PLN 09-DEC-2003
DEFINITION Medicago truncatula clone mth2-53p19, complete sequence.
ACCESSION AC146745
VERSION    AC146745.17 GI:39573788
KEYWORDS   HTG.
SOURCE     Medicago truncatula (barrel medic)
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE  1 (bases 1 to 121952)
AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE     Medicago truncatula BAC Clone mth2-53p19
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 121952)
AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE     Direct Submission
JOURNAL   Submitted (24-SEP-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE  3 (bases 1 to 121952)
AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE     Direct Submission
JOURNAL   Submitted (06-DEC-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE  4 (bases 1 to 121952)
AUTHORS   Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.
TITLE     Direct Submission
JOURNAL   Submitted (09-DEC-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT   On Dec 9, 2003 this sequence version replaced gi:39229182.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
Location/Qualifiers
1..121952
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-53p19"

```

ORIGIN /clone_lib="Medicago truncatula BAC library H2"

Query Match 4.1%; Score 87.8; DB 8; Length 121952;
Best Local Similarity 55.0%; Pred. No. 4.5e-11;
Matches 193; Conservative 0; Mismatches 157; Indels 1; Gaps 1;

QY 1 GGATCCCAACTTTTAGAATGGATCTTAAATTTTACCTTATATAGTTCCAAAGTTAGAAAA 60
|||
Db 94945 GGCTCCATGCTTGAAGACACTTAAATTTTCTTATTCGATCATGAGTGGTATAA 95004
|||
QY 61 TCTTTACCAAGAGCTTTGAGTCCATTCATGACATCCGTGAAC- GGTGTACATGCTCTCCG 119
|||
Db 95005 GTTTTCCAGTGATCTTAATTCATTCACATAGATGTAATCTGGAGTACATTCATCA 95064
|||
QY 120 ATGAGTCACTTGGTTTCATTCGAAAAAGTTCGAAGAGTGCATAGAAATATTGATTTTG 179
|||
Db 95065 AAGTCTCATCTCGTTCATCTTCGAAGATTTCAAAATTTCTTACGCCAATATCTATCTT 95124
|||
QY 180 GATTCTTCACTCGGTGGTGGCTTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAA 239
|||
Db 95125 GTTTCCTTACGGCTAGTTCCTCATGATGATTTTATAGATGTCCTCCAAATCTTTT 95184
|||
QY 240 GCCAATCAAAATGAAATGATGAAATTCATTTTGTCTAATGCACAAAAACAGGGCA 299
|||
Db 95185 GCAGTAACATTCATCACTCTTTCATCTTCTCCATGCTTGTAGTGGCATGATGAAG 95244
|||
QY 300 TTCAATAGCTTTGTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCAT 350
|||
Db 95245 AGTCGAGCTTTGAGTTAGAGTACTTTTTTCTTTATCTTCTTTTGACCAT 95295
|||

RESULT 7

AP004957
LOCUS 132605 bp DNA linear PLN 22-JUL-2003
DEFINITION Lotus corniculatus var. japonicus genomic DNA, chromosome 2,
clone: LjT34H20, TM0134, complete sequence.

ACCESSION AP004957

VERSION AP004957.1

KEYWORDS GI:21907975

SOURCE HTG.

ORGANISM

Lotus corniculatus var. japonicus (Lotus japonicus)

Lotus corniculatus var. japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.

REFERENCE

1 Kaneko, T., Nakamura, Y., Asamizu, E., Kato, T., Sato, S. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. I. Sequence
Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb
Regions of the Genome
Unpublished

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 252-0812, Japan (E-mail: ynk@kazusa.or.jp,
URL: http://www.kazusa.or.jp, Tel: 81-438-52-3935,
Fax: 81-438-52-3934)

FEATURES

source

1. 132605 /organism="Lotus corniculatus var. japonicus"

/mol_type="genomic DNA"

/variety="japonicus"

/db_xref="taxon:34305"

/chromosome="2"

/clone="LjT34H20"

/clone_lib="LjT library"

/note="TAC clone: TM0134-synonym: Lotus japonicus"

ORIGIN

Query Match 4.0%; Score 85.6; DB 8; Length 132605;

Best Local Similarity 50.8%; Pred. No. 1.6e-10;
Matches 268; Conservative 1; Mismatches 240; Indels 19; Gaps 2;

QY 1 GGATCCCAACTTTTAGAATGGATCTTAAATTTTACCTTATATAGTTCCAAAGTTAGAAAA 60
|||
Db 112738 GGATCCAGCTTTTCAGGAAGACCTTAAGAATCTCTATAACATGATCACCAGTTGTGTA 112797
|||
QY 61 TCTTTACCAAGAGCTTTGAGTCCATTCATGATCA-CATCCGTGAACGGTGTACATGCTCTCCG 119
|||
Db 112798 CTTTATTGAGGGGTCTTAATTCAGCAATAGTAATCTGAAATCTGGAGAACATGTCTTCA 112857
|||
QY 120 ATGAGTCACTTGGTTTCATTCGAAAAAGTTCGAAGAGTGCATAGAAATATTGATTTTG 179
|||
Db 112858 ATGAGTCACTGATTCATTCGAAAGATTCATATCTCTGATCAGGACACGCCCTTT 112917
|||
QY 180 GATTCTTCACTCGGTGGTGGCTTCATGAGTGCCTTCATGAGTGCCTCAAGATCCTCCAAATATCAAAA 239
|||
Db 112918 GACTCTTTCACCTTCTTTTCCATCATGGGTCTCTTCAAGGAATCAAGATACCCCTTA 112977
|||
QY 240 GCCAATCAAAATGAAATGATGAAATTCATTTTGTCTAATGCACAAAAACAGGGCA 299
|||
Db 112978 GCATCTCAGATCTTAATCTCTGATTTCTTCAATAGAAATAGCATGAGAAGAATA 113037
|||
QY 300 TTCAATAGCTTTGTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATCGCTCATC 359
|||
Db 113038 GTTCTAGCTTTATGATGCGAGTGAAGTAAAGCTCTTCTGCTCGAGTCTCTGTTTG 113097
|||
QY 360 G-----GAAGAGAAAAATTTTGAATTCATTTTTCGACAAATAGACCAA 401
|||
Db 113098 GAGATCTTCTGTCATCACCATCACTGGAAGCTCGTAACCATCCACTATAATATCCAG 113157
|||
QY 402 AGCTCGAAATCCATGGAATGAGGAAGATCCTCATATGATGATTTTCCATATCATGTAATTC 461
|||
Db 113158 AGATCTGATCAAGGCCCAAGAAAGCTTTCAATCTCTGTTCCAGTACTCAAAATCTT 113217
|||
QY 462 GACTCAATTAACATAGTGTGATGTGAATGAATGACCTTCATGCTCT 509
|||
Db 113218 TGACCATCAACATAGAGGCTTTGCACGTAGACATCTTTTGGCGT 113265
|||

RESULT 8

AC146910/c

LOCUS

DEFINITION

Medicago truncatula clone mth2-59121, WORKING DRAFT SEQUENCE, 2

unordered pieces.

ACCESSION

AC146910

VERSION

AC146910.5

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE

Medicago truncatula (barrel medic)

ORGANISM

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE

1 (bases 1 to 109284)

AUTHORS

Lin S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.

and Roe, B. A.

TITLE

Medicago truncatula BAC Clone mth2-59121

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 109284)

AUTHORS

Lin S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.

and Roe, B. A.

TITLE

Direct Submission

JOURNAL

Submitted (06-DEC-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE

3 (bases 1 to 109284)

AUTHORS

Lin S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.

and Roe, B. A.

TITLE

Direct Submission

JOURNAL

Submitted (06-DEC-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT On Dec 6, 2003 this sequence version replaced gi:38371848.
 ----- Genome Center
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:UOKNOR

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 30237: contig of 30237 bp in length
 * 30238 30337: gap of unknown length
 * 30338 109284: contig of 78947 bp in length.
 FEATURES
 source Location/Qualifiers
 1. 109284
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /db_xref="taxon:3880"
 /clone="mth2-59121"
 /clone_lib="Medicago truncatula BAC library H2"
 ORIGIN
 Query Match 3.9%; Score 83.8; DB 2; Length 109284;
 Best Local Similarity 54.5%; Pred. NO. 4.7e-10;
 Matches 189; Conservative 0; Mismatches 157; Indels 1; Gaps 1;
 QY 7 CAACCTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAAAGTTAGAAATCTTTA 66
 Db 1100 CAAGATCTGGAGACATCTTAAATTTTCTTACTCTTTCATGGATTGAAATCTTTT 1041
 QY 67 CCAAGAGCTTTAGTCCATTGATGACATCCGTGAACGGTG-TACATGTCTCCGATGCAC 125
 Db 1040 TCAAGAGAGCGTAATCTATTGATGATAGTAGTAAACCTTCCATACATTTTCATCTATGTT 981
 QY 126 TCACCTGGTTTCATCGGAAAGTTCGAAAGAGTGATGATGATGATGATGATGATGATGAT 185
 Db 980 TCGGTTTCTTTTCATTTGCAATAGTTCAAACTTTCGAAACCAATGTCAATTTCTTTTCT 921
 QY 186 TTCCTCGTGTGGTGGCTTCATGATGACCTCAAGAGTCCTCCAAATATCAAAAGCCGAA 245
 Db 920 TTTAGTGATGGTGGCTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 861
 QY 246 TCAAAAATTGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 305
 Db 860 TTGCATTTGTTGCTACTCTATCATATTTCTCCCTGCTCAATGCCATGGTTAGAAATAGTTA 801
 QY 306 GCCTTTGTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATTC 352
 Db 800 GCTTTAAATTTAATACATCGCTTACCTTCTTCTTCTGACCAATC 754
 RESULT 9
 AC124959/c 125883 bp DNA linear PLN 05-MAR-2003
 LOCUS Medicago truncatula clone mth2-23111, complete sequence.
 DEFINITION
 AC124959
 ACCESSION
 VERSION AC124959.11 GI:28850030
 HTG
 KEYWORDS
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases 1 to 125883)
 Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
 Cook,D., Kim,D. and Roe,B.A.
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:UOKNOR

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 30237: contig of 30237 bp in length
 * 30238 30337: gap of unknown length
 * 30338 109284: contig of 78947 bp in length.
 FEATURES
 source Location/Qualifiers
 1. 109284
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /db_xref="taxon:3880"
 /clone="mth2-59121"
 /clone_lib="Medicago truncatula BAC library H2"
 ORIGIN
 Query Match 3.9%; Score 83.8; DB 2; Length 109284;
 Best Local Similarity 54.5%; Pred. NO. 4.7e-10;
 Matches 189; Conservative 0; Mismatches 157; Indels 1; Gaps 1;
 QY 7 CAACCTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAAAGTTAGAAATCTTTA 66
 Db 1100 CAAGATCTGGAGACATCTTAAATTTTCTTACTCTTTCATGGATTGAAATCTTTT 1041
 QY 67 CCAAGAGCTTTAGTCCATTGATGACATCCGTGAACGGTG-TACATGTCTCCGATGCAC 125
 Db 1040 TCAAGAGAGCGTAATCTATTGATGATAGTAGTAAACCTTCCATACATTTTCATCTATGTT 981
 QY 126 TCACCTGGTTTCATCGGAAAGTTCGAAAGAGTGATGATGATGATGATGATGATGATGAT 185
 Db 980 TCGGTTTCTTTTCATTTGCAATAGTTCAAACTTTCGAAACCAATGTCAATTTCTTTTCT 921
 QY 186 TTCCTCGTGTGGTGGCTTCATGATGACCTCAAGAGTCCTCCAAATATCAAAAGCCGAA 245
 Db 920 TTTAGTGATGGTGGCTTCATGATGATGATGATGATGATGATGATGATGATGATGAT 861
 QY 246 TCAAAAATTGAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 305
 Db 860 TTGCATTTGTTGCTACTCTATCATATTTCTCCCTGCTCAATGCCATGGTTAGAAATAGTTA 801
 QY 306 GCCTTTGTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATTC 352
 Db 800 GCTTTAAATTTAATACATCGCTTACCTTCTTCTTCTGACCAATC 754

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

2 (bases 1 to 125883)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,

Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (20-JUN-2002) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

3 (bases 1 to 125883)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,

Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (06-NOV-2002) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

4 (bases 1 to 125883)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,

Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (05-MAR-2003) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

On Mar 5, 2003 this sequence version replaced gi:24635962.

Center: Department Of Chemistry And Biochemistry

The University Of Oklahoma

Center code:UOKNOR

Location/Qualifiers

1. 125883

/organism="Medicago truncatula"

/mol_type="genomic DNA"

/db_xref="taxon:3880"

/clone="mth2-23111"

/clone_lib="Medicago truncatula BAC library H2"

Query Match

Best Local Similarity

Matches

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

2 (bases 1 to 125883)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,

Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (20-JUN-2002) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

3 (bases 1 to 125883)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,

Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (06-NOV-2002) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

4 (bases 1 to 125883)

Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,

Cook,D., Kim,D. and Roe,B.A.

Direct Submission

Submitted (05-MAR-2003) Department Of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

On Mar 5, 2003 this sequence version replaced gi:24635962.

Center: Department Of Chemistry And Biochemistry

The University Of Oklahoma

Center code:UOKNOR

Location/Qualifiers

1. 125883

/organism="Medicago truncatula"

/mol_type="genomic DNA"

/db_xref="taxon:3880"

/clone="mth2-23111"

/clone_lib="Medicago truncatula BAC library H2"

Query Match

Best Local Similarity

Matches

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db 4758 GAGCCATCAAGAGTGGTGG 4739

|||||

Query Match 3.9%; Score 83.2; DB 2; Length 141590;
 Best Local Similarity 54.3%; Pred. No. 6.7e-10;
 Matches 189; Conservative 0; Mismatches 158; Indels 1; Gaps 1;

QY 6 CCAACTTTTAGGAATGGATCTTAAATTTTATAGTTTAAAGTTAGAAAAATCTTT 65
 Db 83904 CCAAGATCTTGGGAGACATCTTAAATTTTCTTACTTCTTCAATGATGTTAAATCCTT 83845

QY 66 ACCAAGAGCTTTGAGTCCATTGATGACATCGTGAAACGGTG-TACATGCTCCGATGGA 124
 Db 83844 TTCAAGAGAGCGTAACCTCAITTCATATGATAGTAAACCTTCCATACATTTTCATCTATGGT 83785

QY 125 CTCACCTTGGTTTCAATTCGGAAGCTTCAAGAGTGCATAGAAATATTTGGGATTC 184
 Db 83784 TCGGTTTCTTCAATTCGGAATAGTTTCAAACTTTTCAACACCAATGTCATTTCTTGTTTC 83725

QY 185 TTTCACCTCGTTGGTGGCTTCATGAGTGACCTCAAGAGTCTTCAATGCTCTCAATGCACAAACAGGGCATTCAT 244
 Db 83724 TTTTACGTTGATTTGGTGCTTTCATGATGATGATTAATGATATCCCAAAATTTCTTTGGCATT 83665

QY 245 ATCAAAATTCGAAATGATTTGATTTGATTTTCTCTAATGCACAAACAGGGCATTCAT 304
 Db 83664 TTTCATTTCTTGTACTCTATCATATTTCTTCCTCTCAATGCAATGGTTAGAAATAGTTT 83605

QY 305 AGCCTTTGTGTTTAAAGCAAAACATTTCTTCGATTCATCCCATTC 352
 Db 83604 AGCTTTAAATTAATAGCATGCGCTTACCTTCTTCTTCGACCATTC 83557

RESULT 11
 AC126788 141590 bp DNA linear HTG 11-NOV-2003
 LOCUS Medicago truncatula clone mth2-33121, WORKING DRAFT SEQUENCE, 8
 DEFINITION unordered pieces.
 ACCESSION AC126788
 VERSION AC126788.9 GI:38257184
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE
 1 (bases 1 to 141590)
 Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.
 Medicago truncatula BAC Clone mth2-33121
 Unpublished

REFERENCE
 2 (bases 1 to 141590)
 Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.
 Direct Submission
 Submitted (109-JUL-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA

REFERENCE
 3 (bases 1 to 141590)
 Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D. and Roe, B.A.
 Direct Submission
 Submitted (11-NOV-2003) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA

COMMENT
 On Nov 11, 2003 this sequence version replaced gi:37059850.
 Center: Genome Center
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:UOKNOR

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 4380: contig of 4380 bp in length
 * 4381 4480: gap of unknown length
 * 4481 14258: contig of 9778 bp in length
 * 14259 14358: gap of unknown length
 * 14359 20669: contig of 6311 bp in length
 * 20670 20769: gap of unknown length
 * 20770 25083: contig of 4314 bp in length
 * 25084 25183: gap of unknown length
 * 25184 41406: contig of 16223 bp in length
 * 41407 41506: gap of unknown length
 * 41507 60351: contig of 18845 bp in length
 * 60352 60451: gap of unknown length
 * 60452 87168: contig of 26716 bp in length
 * 87168 87267: gap of unknown length
 * 87268 141590: contig of 54323 bp in length.

Location/Qualifiers
 1. .141590
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /db_xref="taxon:3880"
 /clone="mth2-33121"
 /clone_lib="Medicago truncatula BAC library H2"

FEATURES
 source

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 4380: contig of 4380 bp in length
* 4381 4480: gap of unknown length
* 4481 14258: contig of 9778 bp in length
* 14259 14358: gap of unknown length
* 14359 20669: contig of 6311 bp in length
* 20670 25083: gap of unknown length
* 25084 25183: gap of unknown length
* 25184 41406: contig of 16223 bp in length
* 41407 41506: gap of unknown length
* 41507 60351: contig of 18845 bp in length
* 60352 60451: gap of unknown length
* 60452 87167: contig of 26716 bp in length
* 87168 87267: gap of unknown length
* 87268 141590: contig of 54323 bp in length.
FEATURES
    source
    1..141590
        /organism="Medicago truncatula"
        /mol_type="genomic DNA"
        /db_xref="taxon:3880"
        /clone="mth2-33121"
        /clone_lib="Medicago truncatula BAC library H2"

```

ORIGIN

```

Query Match      3.8%; Score 83; DB 2; Length 141590;
Best Local Similarity 54.1%; Pred. No. 7.5e-10;
Matches 190; Conservative 0; Mismatches 160; Indels 1; Gaps 1;

Qy 1 GGATCCCAACTTTAGGAATGGATCTTAAATTTTAGTTATAGTTCAAAGTTAGAAAA 60
Db 40054 GGCTCCATGTCCTTGAAGACATCTTAAGATTTTCTTATCCGATCATGATGAT 40113
Qy 61 TCTTTACCAAGAGCTTTGAGTCCATTCATGATCATCCGTTGAAC-GGTGTACATGCTCCG 119
Db 40114 GTTTTCCAGTGAATCTTAATCAATCACAATAGATGATCTCGAGTACATTTATCA 40173
Qy 120 ATGACATCATTGGTTTCATTCGAAAGTTTGAAGAGTGCATPAAGAATATTCATTTG 179
Db 40174 ATGTTCTCATCTTCGTTCAATTCAGAGATTTCAAATTTCTACGCCAATGCTCATTT 40233
Qy 180 GATTTCTTCACTCGTTGGTGCCTTCATGATGACCTCAGAGTCCTCCAAATATCAAA 239
Db 40234 GTTTCTTTTACATGGCTAGTTCCTTCATGATGAATTTTATAGAGTGTCCCAACACTTCTTT 40293
Qy 240 GCCGAATCACAAATGAAATGTGATTGAATTCATTTTGTCTAATGCACAAACACAGGCA 299
Db 40294 GCAGTCACACATTCATCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 40353
Qy 300 TTCATAGCTTTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCAT 350
Db 40354 AGTCGAGCTTTAGAGTTTAGAAGTACTTCTCTTTATCTTCTTCTTCTTCTTCTTCTT 40404

```

RESULT 12

```

AF411804
LOCUS              100810 bp      DNA      linear      PLN 25-JUL-2002
DEFINITION        Lycopersicon esculentum BAC clone Clemson_ID 47113, complete
sequence.
ACCESSION         AF411804
VERSION           AF411804.1
KEYWORDS          GI:15987769
SOURCE            Lycopersicon esculentum (tomato)
ORGANISM          Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE         1 (bases 1 to 100810)
AUTHORS           Van Der Hoeven,R., Ronning,C., Giovannoni,J., Martin,G. and

```

TITLE

```

JOURNAL            Tanksley,S.
PUBMED            12119366
REFERENCE          2 (bases 1 to 100810)
AUTHORS           van der Hoeven,R.S. and Tanksley,S.D.
TITLE             Direct Submission
JOURNAL           Submitted (20-AUG-2001) Plant Breeding, Cornell University, 252
Emerson Hall, Ithaca, NY 14850, USA
FEATURES
    source
    1..100810
        /organism="Lycopersicon esculentum"
        /mol_type="genomic DNA"
        /db_xref="taxon:4081"
        /clone="Clemson_ID 47113 BAC"

```

ORIGIN

```

Query Match      3.8%; Score 81.2; DB 8; Length 100810;
Best Local Similarity 52.0%; Pred. No. 2.1e-09;
Matches 205; Conservative 0; Mismatches 188; Indels 1; Gaps 1;

Qy 4 TCCCACTTTTAGGAATGGATCTTAAATTTTAGTTATAGTTCAAAGTTAGAAAAATCT 63
Db 89844 TCCCAAGACCTAGGAAAAATCCTAAGTACTTTGTCAACCTGTTCTCAGTGGTATATACT 89903
Qy 64 TTACCAGAGCTTTGAGTCCATTCATTCATCGATCGTGAAC-GGTGTACATGTCCTCCGATG 122
Db 89904 TTACCAATGAGATTTAGTTTATCAGATGCTGAATCTAGTAATCATATATCTTGAAGA 89963
Qy 123 GACTCATCTTGGTTTCATTCGAAAGTTTGAAGAGTGCATGAATATTCATTTTGGAT 182
Db 89964 GATTTCCGATTTCACTTAAAGCTTCATATTCGACACAACTAGCAATCTTGAAT 90023
Qy 183 TCTTTCACTCGTTGGTGCCTTCATGATGACCTCAAGAGTCCTCCAAATATCAAAAGCC 242
Db 90024 TTTCCGACTTTGAGTTGACCTTCATGAGAGTTTGTAAAGTATCCCATATTTGTTTAGCA 90083
Qy 243 GAATCAAAATGAAATGATGATTAATTTGTTGTTCTAATGCACAAACAGGCGATTC 302
Db 90084 GAGGTACAACTTGAGATACGTTTAACTCATCGGTCCTAGTCCACATCAAGAATGCAC 90143
Qy 303 ATAGCCTTTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATTCGCTCATCGGA 362
Db 90144 TTAGCTTTGCATCTTCTTCTAGCATTTTATATCCGCTTCATATATGATCTCGAGGT 90203
Qy 363 AGAGAAATTTTGAATCCATTTTCGACAATAG 396
Db 90204 TTAGGTACTTTATTTCTTCTCACTATCAGTAATAG 90237

```

RESULT 13

```

AC087797
LOCUS              122497 bp      DNA      linear      PLN 29-AUG-2001
DEFINITION        Oryza sativa chromosome 3 BAC OSUNBB0022E02 genomic sequence,
complete sequence.
ACCESSION         AC087797
VERSION           AC087797.5
KEYWORDS          GI:13699786
SOURCE            Oryza sativa (japonica cultivar-group)
ORGANISM          Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE         1 (bases 1 to 122497)
AUTHORS           Buell,C.R., Yuan,Q., Ouyang,S., Moffat,K.S., Hill,J.N.,
Gansberger,K., Brenner,M., Burgess,S., Hance,M., Shvartsbeyn,M.,
Tsitin,T., Riggs,F., Hsiao,J., Zismann,V., Blunt,S., Pai,G.,
VanAken,S.E., Utterback,T.R., Feldblyum,T.V., Quackenbush,J.,
Salzberg,S.L., White,O. and Fraser,C.M.
SOURCE            Oryza sativa chromosome 3 BAC OSUNBB0022E02 genomic sequence
Unpublished

```



```
TTTKDQCQIEGPFETHSILEGTHLCELDWCWSQSMKILGESSRKLTKDLGV"
complement(18822..18858)
/rpt_family="(CAGC)n"
complement(19612..19640)
/rpt_family="AT rich"
complement(20459..20487)
/rpt_family="AT rich"
21271..21576
/gene="OSUNBB0022E02.13"
/notes="predicted by fgenesh"
<21271..>21576
/gene="OSUNBB0022E02.13"
21271..21576
/gene="OSUNBB0022E02.13"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK50120.1"
/db_xref="GI:13937289"
/translation="MTTRTVLLSEYSDSHNSGRNGAVGVGHGVGAGGRGLHGYI
ACAVLLASTPYAPLHTQGELOHGVWTKDCEWRQCAPWDVDYEFTSNDASVPNS"
complement(22507..22529)
/rpt_family="(TA)n"
complement(22544..22588)
/rpt_family="AT rich"
complement(23124..23217)
/rpt_family="(GGAGAA)n"
complement(23165..25551)
/notes="predicted by fgenesh"
complement(join(23165..23588,23765..23917,25379..>25551))
/gene="OSUNBB0022E02.14"
complement(join(23165..23588,23765..23917,25379..25551))
/gene="OSUNBB0022E02.14"
/codon_start=1
/product="hypothetical protein"

Query Match      3.8%; Score 81; DB 8; Length 122497;
Best local similarity 48.5%; Pred. No. 2.4e-09;
Matches 314; Conservative 1; Mismatches 321; Indels 11; Gaps 3;

QY 4 TCCCAACTTTTGGAGTGGATCTTTAAATTTTGTATTAAGTTCAAAAGTTAGAAAATCT 63
DB 62908 TCATATCTTCGGGTAGGCTCTAAGCATCTTTTGGGCAACCTCAAGATCGGTAGTTT 62967
QY 64 TTACCAAGAGTTTGAAGTCATTCATGACATCCGTGAAC-GGTGTACATCTCCCATG 122
DB 62968 GCCCAAGCCCTTGAGATCATTTACAAATGACATGTAGCCCTCCCATATCATATCACA 63027
QY 123 GACTCACTTGTGTTTCATTCGAAAGTTGCAAGAGTGCAAGAAATATTCATTTTGGAT 182
DB 63028 CTCTCATGAGCAACATGGAGAAATGTCATATTTGGATCTTGAGAAATGAGCTTGGCA 63087
QY 183 TCTTTCACTCGGTTGGTCCCTTCATGATGACCTCAAGAGTCCTCCAAATATCAAAAGCC 242
DB 63088 TCCCTGTACTCACCTGTATCCCTCATGATCTCCGCAACTGTTCCAAATCTCATATGCG 63147
QY 243 GAATCACAAATGAAATGTGATGAATTCATTTTGTCTATGACAAACAGGCGATTC 302
DB 63148 GTCTCAAGGTTGCTCACTCTATCGAATCTCTTGGGTCAAAGAGTTGAAACAGGCAATC 63207
QY 303 ATAGCCTTTGTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATTCCTCATCGGA 362
DB 63208 ATGGC-----TTGGGCATTGAGTTGGAGTTGGGTTGATCAATCTCCGTCAAGGT 63258
QY 363 AGAGAAATTTTGAATTCATTTTTCGACAAATAGACCAAGCTCGAATTCATCGAATG 422
DB 63259 GTGCGGTTGATGGCAAGCCCTACATCCCAATATCTCCAAATATGGAAGCTCATAGCT 63318
QY 423 AGGAAGATCCTCATATGATTTTCCAAATACATGTAATTCGACTCATTAACATAGGTGGA 482
DB 63319 AGGTGAGTAGACATTTTAAATTTTCCAGTGGAGTAGTTTGTGCCATTAACATGGAGCC 63378
QY 483 TGTGTAATGAAATGACCCCTCATGCSCTATC-TCTCTTGGGTATTAAACCAAAATATGAG 541
```

```
Db 63379 TTCCCTACATGGTTCCACTCGTTTCGACATCTTCTCTAGCGGTGAAGCCCAATCAAGA 63438
QY 542 TGAGCCCTTGTCTGTATACCAATTTGTTAGATCAGATGGCACTAAGAGAGGGGGAGTG 601
Db 63439 GAGACCAAGCTCTGATACCACTTTGTAGATCGAGATGTCGACTAGAGGGGGTGAATAGG 63498
QY 602 AATTAGTCGATCGATTAAACCTTATAAGTTTAAATAATGAATTCGTA 648
Db 63499 CGATTTAAACTAAATAACACCTAATCAAACTAACCTAAGTTGCTA 63545
```

```
RESULT 14
AC144740 157436 bp DNA linear HTG 15-MAY-2003
LOCUS Oryza sativa (japonica cultivar-group) chromosome 5 clone
DEFINITION OSUNBa0075A10, *** SEQUENCING IN PROGRESS ***, 9 ordered pieces.
ACCESSION AC144740
VERSION AC144740.1 GI:30725895
KEYWORDS HTG; HTGS PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 157436)
AUTHORS Chow,T.-Y., Hsing,Y.-I.C., Chen,C.-S., Chen,H.-H., Liu,S.-M.,
Chao,Y.-T., Lee,P.-F., Chang,S.-J., Chen,H.-C., Chen,S.-K.,
Chen,T.-R., Chen,Y.-L., Cheng,C.-H., Chung,C.-I., Han,S.-Y.,
Hsiao,S.-H., Hsiung,J.-N., Hsu,C.-H., Kau,P.-I., Lee,M.-C.,
Liu,H.-L., Li,Y.-F., Lin,S.-J., Lin,Y.-C., Wu,S.-W., Yu,C.-Y.,
Yu,S.-W., Wu,H.-P. and Shaw,J.-F.
```

Oryza sativa BAC OSUNBa0075A10 genomic sequence

```
Unpublished
2 (bases 1 to 157436)
Chow,T.-Y. and Hsing,Y.-I.C.
Direct Submission
Submitted (15-MAY-2003) Institute of Botany, Academia Sinica, 128,
Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
```

```
* 1 5300: contig of 5300 bp in length
* 5301 5400: gap of unknown length
* 5401 17970: contig of 12570 bp in length
* 17971 18070: gap of unknown length
* 18071 37874: contig of 19804 bp in length
* 37875 37974: gap of unknown length
* 37975 62822: contig of 24848 bp in length
* 62823 62922: gap of unknown length
* 62923 64710: contig of 1788 bp in length
* 64711 64810: gap of unknown length
* 64811 71025: contig of 6215 bp in length
* 71026 71126 118711: contig of 47586 bp in length
* 118712 118811: gap of unknown length
* 118812 121761: contig of 2950 bp in length
* 121762 121861: gap of unknown length
* 121862 157436: contig of 35575 bp in length.
```

Location/Qualifiers

```
1. 157436
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="5"
/clones="OSUNBa0075A10"
```


ORIGIN

Query Match 3.8%; Score 81; DB 2; Length 157436;
Best Local Similarity 48.5%; Pred. No. 2.4e-09;
Matches 314; Conservative 1; Mismatches 321; Indels 11; Gaps 3;
QY 4 TCCCACTTTTGGAAATGGATCTTAAATTTTGGTTATTAAGTTCAAAAGTTAGAAAAATCT 63
Db 96233 TCATACCTCTCCGGTAGAGGCTAAGCATCTTTGGGCAACCTCAAGATCATGATGATTT 96292
QY 64 TTACCAAGAGCTTTGAGTCCATTTGATGACATCCGTAAGC-GGTGTACATCTCCCGATG 122
Db 96293 GCCCAAGCCCTTTGAGATCATTCACATGATCATAGCTCCCATCATGCTATTACA 96352
QY 123 GACTCACTGTGTTTCATTCGAAAGAGTTCGAAAGATGTCATAAGAAATATTGATTTGGAT 182
Db 96353 CTCTCATGAGCAACATGGAGAAATGCTCATATTGGATCTTTAGGAATGAAGCTTGCCA 96412
QY 183 TCTTTCACTCGGTGGTGGCTTCAATGATGACCTCAAGAGTCCCTCCAAATATCAAAAGCC 242
Db 96413 TCCTTGTACTCGTGTGGCCCTCATGGATCTCCGCCAATCTTCCAAATCTCATATGCG 96472
QY 243 GAATCACAAAATGAAATGATGATGAAATCATTTTGTCTAATGACACAAACAGGGCATTC 302
Db 96473 GTCTTAAGTGTGCTCACTCATCGAACTCTCTTGGCTCAAAGAGTTGAACAAAGCATTC 96532
QY 303 ATAGCCTTGTGTTTAAAGCAAAACATCTCTCCGATTCATCCATTCCTCATCGGA 362
Db 96533 ATAGCTTGGCATGATGTTGGAGTGTGGTGTGATCAATCTCCGTCAAAGGGGCTCGGTG 96592
QY 363 AGAGAAATTTTGAATCCATTTTCGACATGACCAAGCTCGAAATCCATGGAATG 422
Db 96593 ATAGTAAAGCCTACA-----TCCACATATCCCAATATGGAACCTCATAGCTTTG 96643
QY 423 AGAAGATCTCATATGAGTTTCCAAATACATGATTCGACTCATTAACATAGTGGGA 482
Db 96644 AGGTGATGACATATTTAAATTTTCCAAAGTGGAGTGTGTGCTTAAACATGAGGAGCC 96703
QY 483 TGTGTATGAATGACCTCATGCTATC-TCTCTGGGTATTAACCAATATGAGAG 541
Db 96704 TTCCTTACATGTTTCACTCGTTCGACATCTCTCTAGGGCGTAAAGCCCAATCAAGA 96763
QY 542 TGAGCCTTGTCTGATACCAATTTGTTAGGATGACAGTGGCACTAAGAGAGGGGGAGTG 601
Db 96764 GAGCAACAGCTCTGATACCACTTGTAGATCGAGATCGACTAGAGGGGGTGAATAGG 96823
QY 602 AATTAGTGCAGTGGATTAACATTAAGTTTAAATGAATTCGTA 648
Db 96824 CGATTTAAACTAAATGACACCTAATCAAACTAATCACTAAGTTACTA 96870

RESULT 15
AP003198/c

LOCUS AP003198 135301 bp DNA linear PLN 27-NOV-2003
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
BAC clone: B1039D07.
ACCESSION AP003198 BA000010
VERSION AP003198.3 GI:15408621
KEYWORDS Oryza sativa (japonica cultivar-group)
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T.,
Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nakamura, Y.,
Antonio, B. A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K.,
Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K.,
Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y.,
Itonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Itoh, S., Itoh, T.,
Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S.,
Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T.,
Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M.,
Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M.,
Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T.,
Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K.,
Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K.,
Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J. H., Kim, H. I., Eun, M. Y.,
Yano, M., Jiang, J. and Gojobori, T.
The genome sequence and structure of rice chromosome 1
Nature 420 (6913), 312-316 (2002)
22337376
12447438
2 (bases 1 to 135301)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Aug 31, 2001 this sequence version replaced gi:14522858.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLAST2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from M13rev to -21M13 of the BAC
clone. This sequence of B1039D07 clone has an overlap with P0493G01
(DBJ: AP002914) clone at the position 1 to 55,572 and with
B1045D11 (DBJ: AP003199) at the position 63,818 to 135,301 of 5'
end. The sequence of this clone starts at the position 11,2015 of
P0493G01 and ends at the position 71,483 of B1045D11. Detailed
information on overlap and assembly quality together with
annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES
Location/Qualifiers
1..135301
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="1"
/clone="B1039D07"
complement(join(1884..1995,2143..2186))
/gene="B1039D07.1"
complement(join(1884..1995,2143..2186))
/gene="B1039D07.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64039.1"
/db_xref="GI:15408622"
/translation="MTAGGEEGEGVRSSWFTTSGLIYEEVVTDRMDTAEGGEERGE
GYRRQWD"
complement(join(4700..5160,5990..6038,6142..6291))
/gene="B1039D07.2"
complement(join(4700..5160,5990..6038,6142..6291))
/gene="B1039D07.2"
/codon_start=1
/product="putative glutathione S-transferase"
/protein_id="BAB64040.1"
/db_xref="GI:15408623"
/translation="WTPVKVFGPAQSTNVARVLLCLCEVGAEYEVNVDFVTMEHKSP
EHLKRNPFQIPAFQDGLDLYFESRAIGKYILRKYTKREADLLREGNREAAVVDVMT

Db 94628 TTCCCTACATGGTTACCTCGTTGACATCTTCTCTAGTGGTGAAGCCCAATCAAGA 94569
 QY 542 TGAGCCTTGCTCTGATACCAATTGTTAGGATCAGAGTGGCACTAAGAGAGGGGGGA 598
 Db 94568 GAGACCAAGCTCTGATACCACTTGTAAAGATCGAGATGTCGACTAGAGGGGGGTGAA 94512

Search completed: September 18, 2004, 23:51:53
 Job time : 8127.47 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 14:18:43 ; Search time 761.294 Seconds
(without alignments)
12030.986 Million cell updates/sec

Title: US-09-892-635A-44

Perfect score: 2156

Sequence: 1 ggatcccaacttttaggaat.....agctctcattcactatgg 2156

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_29Jan04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002s.*

7: geneseqn2003as.*

8: geneseqn2003bs.*

9: geneseqn2003cs.*

10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2120.8	98.4	4924	2 AAX25614	Banana ri
2	75	3.5	809	2 AAV28690	Ripening
3	62.8	2.9	2417	2 AAV22222	SIRE-1 CD
4	54	2.5	110000	3 AAF22305_02	Continuation (3 of
5	53.4	2.5	110000	3 AAF22305_03	Continuation (4 of
6	52.6	2.4	515	2 AAO81475	HMG2 prom
7	51.2	2.4	2000	7 ADA71938	Rice gene
8	46.8	2.2	16439	6 ABL12886	Human imm
9	46.2	2.1	2000	7 ADA71938	Rice gene
10	45	2.1	742	2 AAV28672	Ripening
11	44	2.0	1186	2 AAX25608	Banana ri
12	41.2	1.9	1722	7 ACF73552	Straphyloc
13	41.2	1.9	2881	2 AAV74565	Human imm
14	41.2	1.9	5879	6 ABL12229	Human met
15	41.2	1.9	8197	6 ABL170542	Chemical
16	41.2	1.9	8197	6 ABL170542	Chemical
17	41	1.9	3825	6 AAX13945	Clostridi
18	41	1.9	3825	6 AAX13945	Clostridi
19	41	1.9	6022	4 AAS46662	Tumour su
20	40.8	1.9	7351	6 ABL32028	Human imm
21	40.8	1.9	18967	4 ABL23560	Drosophil
22	40.6	1.9	726	6 ABQ68510	Listeria
23	40.6	1.9	1337	6 ABQ70519	Listeria

RESULT 1

AAX25614

ID AAX25614 standard; DNA; 4924 BP.

XX

AC AAX25614;

DT 02-AUG-1999 (first entry)

XX

DE Banana ripening fruit chitinase DNA.

XX

DE Banana; fruit ripening; differential expression; fruit development;

XX transgenic plant; chitinase; ss.

XX

OS Musa acuminata.

XX

PN WO9915668-A2.

XX

PD 01-APR-1999.

XX

PF 23-SEP-1998; 98WO-US003343.

XX

PR 25-SEP-1997; 97US-0060062P.

XX

PA (BOYC-) BOYCE THOMPSON INST PLANT RES.

XX

PI May G, Clendennen S;

XX

DR WPI: 1999-244425/20.

XX

DR P-PSDB; AAY05845, AAY05846, AAY05847.

XX

PT New isolated banana DNA molecules.

XX

PS Claim 5; Fig 17A-G; 143pp; English.

XX

CC This is the nucleotide sequence of a chitinase DNA molecule. The invention provides isolated DNA molecules which are differentially expressed during banana fruit development, and the protein products of these genes. The DNA is selected from a group comprising starch synthase, chitinase, endochitinase, beta-1,3-glucanase, thaumatin-like protein, ascorbate peroxidase, metallothionein, lectin and senescence-related protein. The regulatory elements of the genes can be used to produce chimeric genes for transformation of plants to provide controlled expression of heterologous DNA during fruit development, or in response to exogenous developmental signals, such as ethylene signals. The heterologous protein, e.g. a therapeutic protein, can be isolated from the fruit or consumed directly in the transformed fruit. The identity of

ALIGNMENTS

C 24	40.6	1.9	6216	6	ABK39932	Human che
C 25	40.6	1.9	6216	6	ABL70139	Chemical
C 26	40.4	1.9	6161	6	ABL32623	Human imm
C 27	40.4	1.9	6237	6	ABL32925	Human imm
C 28	40.2	1.9	15399	6	ABL33515	Human imm
C 29	40	1.9	9483	6	ABL32377	Human imm
C 30	40	1.9	9483	6	ABL70516	Chemical
C 31	40	1.9	9483	6	AAS61092	Human gen
C 32	39.6	1.8	1722	7	ACA19762	Prokaryot
C 33	39.6	1.8	83391	6	ABQ67093	Human ang
C 34	39.4	1.8	9504	6	ABK28407	DNA trans
C 35	39.2	1.8	2150	6	ABK85746	DNA encod
C 36	39.2	1.8	18679	6	ABN80300	Human che
C 37	39	1.8	5815	6	ABL33234	Human imm
C 38	39	1.8	5987	6	ABL33562	Human imm
C 39	39	1.8	5987	6	ABQ67101	Human ang
C 40	39	1.8	5987	9	ADB54307	Pretrate
C 41	39	1.8	5987	9	ADB54179	Pretrate
C 42	39	1.8	6089	4	AAS46579	Tumour su
C 43	39	1.8	6089	6	ABL33836	Human imm
C 44	39	1.8	12025	6	ABL33299	Human imm
C 45	38.8	1.8	3398	4	ABL18844	Drosophil

CC the present sequence is not given. The entire chitinase DNA sequence has
CC been decoded in all 3 reading frames to provide the amino acid sequences
CC given in AAY05845-47

Sequence 4924 BP: 1304 A; 1154 C; 1037 G; 1425 T; 0 U; 4 Other;

Query Match	98.4%	Score 2120.8	DB 2	Length 4924
Best Local Similarity	99.6%	Pred. No. 0		
Matches 2149	Conservative 0	Mismatches 3	Indels 5	Gaps 2
QY 1	GGATCCCAACATTTTAGGAATGGATCTTAAAAATTTTAGTATTAAAGTTCAAAGTTAGAAAAA	60		
DB 1	GGATCCCAACATTTTAGGAATGGATCTTAAAAATTTTAGTATTAAAGTTCAAAGTTAGAAAAA	60		
QY 61	TCTTTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAAACGGTGACATGCTCCGA	120		
DB 61	TCTTTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAAACGGTGACATGCTCCGA	120		
QY 121	TGGACTCACTTGGTTTCATTCCGAAAAAGTTGAAAGAGTGCATAAGAATATTGATTTGG	180		
DB 121	TGGACTCACTTGGTTTCATTCCGAAAAAGTTGAAAGAGTGCATAAGAATATTGATTTGG	180		
QY 181	ATTCTTTCACTCGGTTGGTGCCCTTCATGAGTACCTCAAGAGTCTCCAAATATCAAAG	240		
DB 181	ATTCTTTCACTCGGTTGGTGCCCTTCATGAGTACCTCAAGAGTCTCCAAATATCAAAG	240		
QY 241	CCGAATCAAAATGAAATGTGATTTCAATTTTGTCTAAATGCACAAAACACAGGCAT	300		
DB 241	CCGAATCAAAATGAAATGTGATTTCAATTTTGTCTAAATGCACAAAACACAGGCAT	300		
QY 301	TCATAGCCTTTGGTTTAAAGCAAAAACAATCTTCTCCGATTCATCCCATTCGCTCATCG	360		
DB 301	TCATAGCCTTTGGTTTAAAGCAAAAACAATCTTCTCCGATTCATCCCATTCGCTCATCG	360		
QY 361	GAAGAGAAAATTTTGAATTCATTTTGCACAAATAGACCAAAGCTCGAAATCCATGGAAA	420		
DB 361	GAAGAGAAAATTTTGAATTCATTTTGCACAAATAGACCAAAGCTCGAAATCCATGGAAA	420		
QY 421	TGAGGAAGATCCTCATATGAGTTTCCAAATCATGTAAATTCGACTATTAAACATAGGTG	480		
DB 421	TGAGGAAGATCCTCATATGAGTTTCCAAATCATGTAAATTCGACTATTAAACATAGGTG	480		
QY 481	GATGTGTAATGAAATGACCTCATGCSCTATCTCTCTGGGTATTAAACCAAAATATGAGA	540		
DB 481	GATGTGTAATGAAATGACCTCATGCSCTATCTCTCTGGGTATTAAACCAAAATATGAGA	540		
QY 541	GTGAGCCTTGTCTGTATACCAATTTGTAGGATCAGAGTGGCATAGAGAGGGGGGAGT	600		
DB 541	GTGAGCCTTGTCTGTATACCAATTTGTAGGATCAGAGTGGCATAGAGAGGGGGGAGA	600		
QY 601	GAAATTAGTCAGTGGATTTAAACTTATAGTTTAAAAATGAATTCGTAAATACGAGAAGA	660		
DB 601	GAAATTAGTCAGTGGATTTAAACTTATAGTTTAAAAATGAATTCGTAAATACGAGAAGA	660		
QY 661	TTTTCGTTTTTAATAGTAACTTGAGTAGATGAAAACCAAAAAGTTAAACAGTAGTGTAAATAAC	720		
DB 661	TTTTCGTTTTTAATAGTAACTTGAGTAGATGAAAACCAAAAAGTTAAACAGTAGTGTAAATAAC	720		
QY 721	AAATTCGGGAAGTAGAACTCACATTTCAAGGACATACCAATTTAAAGTGGTTCGGT	780		
DB 721	AAATTCGGGAAGTAGAACTCACATTTCAAGGACATACCAATTTAAAGTGGTTCGGT	780		
QY 781	CAAAATGACCTCATATCCACTTGTGAAGCCCTTCTTCGAAGAGGCTCCCAACTCCCATAGC	840		
DB 781	CAAAATGACCTCATATCCACTTGTGAAGCCCTTCTTCGAAGAGGCTCCCAACTCCCATAGC	840		
QY 841	AAATCACTTTGAAGGGGAAGACAAATACCTCTCTTACNACCTTTTACAAATGGTTCATAC	900		
DB 841	AAATCACTTTGAAGGGGAAGACAAATACCTCTCTTACNACCTTTTACAAATGGTTCATAC	900		
QY 901	TCTTTACAAATTTTCAACGAGAAAGAGGGTGAACATCAAGCAATTTGAAAAACAGACT	960		
DB 901	TCTTTACAAATTTTCAACGAGAAAGAGGGTGAACATCAAGCAATTTGAAAAACAGACT	960		

QY 2041 TCCTACAT-ACCATTGCGCATGGCTTTGTTGATGACAGACACCAAGCTTGCTTTGG 2099
 Db |||||||
 2041 TCCTACATACCATGTCGATGGCTTTGTTGATGACAGACACCAAGCTTGCTTTGG 2100
 QY 2100 TTGTGCTTAACAGAGAGAGAGAGAGAGACCGATAGCTCTCTCATTCATCTATGG 2156
 Db |||||||
 2101 TTGTGCTTAAC-AGAGAGAGAGAGAGACCGATAGCTCTCTCATTCATCTATGG 2153

RESULT 2
 AAV28690
 ID AAV28690 standard; cDNA; 809 BP.
 AC AAV28690;
 XX
 DT 29-JUL-1998 (first entry)
 XX
 DE Ripening banana pulp cDNA clone U-D86 SEQ ID NO:48.
 XX
 KW Banana; ripening; pulp; Musa acuminata cv. Grand Nain; fruit;
 KW genetic control; tissue senescence; ss.
 XX
 OS Musa acuminata.
 XX
 PN WO9811228-A2.
 XX
 PD 19-MAR-1998.
 XX
 PF 08-SEP-1997; 97WO-GB002424.
 XX
 PR 10-SEP-1996; 96GB-00018862.
 PR 25-APR-1997; 97GB-00008366.
 XX
 PA (ZENE) ZENECA LTD.
 XX
 PI Seymour GB, Bird CR, Medina-Suarez RDJ;
 XX
 DR WPI; 1998-207389/18.
 XX
 PT Modulation of ripening or tissue senescence in bananas - comprises use of
 PT DNA isolated from ripening banana pulp to produce genetically modified
 PT fruit.
 XX
 PS Claim 1; Page 48-49; 72pp; English.
 XX

CC The present sequence represents a cDNA clone isolated from ripening
 CC banana pulp. 57 clones were isolated and are given in AAV28643 to
 CC AAV28699. The cDNA clone sequences can be used in a method of modulating
 CC ripening or tissue senescence process in plants of the genus Musa. The
 CC method comprises: (a) inserting into the plant material at least 1 of the
 CC 57 sequences (as above); (b) regenerating the plant material; and (c)
 CC selecting from the transformed regenerants, plants with modulated
 CC ripening or tissue senescence characteristics. Also described in the
 CC present invention are: (1) plants, their progeny, seed and material
 CC obtained from the plants, produced by the above method; (2) a vector
 CC functional in plants comprising a promoter region which is operably in
 CC plant cells, a polynucleotide sequence as defined above, and a
 CC transcription termination sequence; and (3) a method of controlling plant
 CC pathogens comprising the application of anti-pathogenic agent to plants
 CC of (1)
 XX
 SQ Sequence 809 BP; 191 A; 233 C; 197 G; 180 T; 0 U; 8 Other;
 Query Match 3.5%; Score 75; DB 2; Length 809;
 Best Local Similarity 100.0%; Pred. No. 1.4e-10;
 Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2082 CCACAGCTTGCTTTGGTTGGCTTAACAGAGAGAGAGAGAGACACCGATAGCCT 2141
 Db |||||||
 9 CCACAGCTTGCTTTGGTTGGCTTAACAGAGAGAGAGAGAGACACCGATAGCCT 68
 QY 2142 CTCATTCATCTATGG 2156
 Db |||||||

Db 69 CCTCATTCACATATGG 83
 RESULT 3
 AAV22222/c
 ID AAV22222 standard; cDNA; 2417 BP.
 XX
 AC AAV22222;
 XX
 DT 23-JUL-1998 (first entry)
 XX
 DE SIRE-1 cDNA from lambda gt11 soybean cDNA library.
 XX
 KW SIRE-1; Gm776 fragment; soybean; retrotransposon; copia; Ty1; ss.
 XX
 OS Glycine max.
 XX
 PN WO9809505-A1.
 XX
 PD 12-MAR-1998.
 XX
 PF 25-AUG-1997; 97WO-US014802.
 XX
 PR 09-SEP-1996; 96US-0025853P.
 XX
 PA (LOYO) UNIV LOYOLA CHICAGO.
 XX
 PI Laten HM;
 XX
 DR WPI; 1998-193245/17.
 XX
 PT New plant retrovirus sequence, SIRE-1 - useful for, e.g. genetic
 PT engineering in plants.
 XX
 PS Claim 1; Fig 6; 199pp; English.
 XX

CC This sequence represents the SIRE-1 cDNA sequence, isolated from a
 CC lambda gt11 soybean cDNA library, and is a polynucleotide of the
 CC invention. SIRE-1 is a member of the copia/Ty1 retrotransposon family.
 CC SIRE-1 can be used for genetic engineering in plants, e.g. for conferring
 CC resistance to disease, insect infestation, herbicide tolerance, enhance
 CC N2 fixation, or enhanced vigour
 XX
 SQ Sequence 2417 BP; 807 A; 460 C; 596 G; 554 T; 0 U; 0 Other;
 Query Match 2.9%; Score 62.8; DB 2; Length 2417;
 Best Local Similarity 50.6%; Pred. No. 7.6e-07;
 Matches 177; Conservative 0; Mismatches 172; Indels 1; Gaps 1;

QY 4 TCCCAACITTTAGGAATGGATCTTAAATTTTAGTTATAGTTCAAGTTAGAAAATCT 63
 Db |||||||
 786 TCAAATCTCTTAGGCAAGGATCTGAGATCTTTCTCACCAGCTTTTCATCAGTCATCTT 727
 QY 64 TTACCAAGAGCTTTGAGTCCATTTGATGACATC-CGTGAAACGGTGACATGCTCCGATG 122
 Db |||||||
 726 TCTCCCAAGGCACTGCAAGCATTTGCAATTTCAAGATGTTCTATGTAAGTCATGAATA 667
 QY 123 GACTCACTTGGTTTCATTTCGGAAAAAGTGTGCAAGAGTGCATAGAATATTTGATTTGGAT 182
 Db |||||||
 666 CACTCTCTCTCTCTCATCTTCAGATTTTCGAATTTTGTGGCCCAATAGTTGCAATCTGGAC 607
 QY 183 TCTTTCACTCGTGTGGTTCCTTCATGAGTGACCTCAAGAGTCTTCCAAATATCAAAAGCC 242
 Db |||||||
 606 ATCTTCACCTTTGGAGGTTCTCTTCATGAGTGGTTTTTCAGGATCTCCCATGCTCTTGGCC 547
 QY 243 GAATCACAATAATGAATGTGATTTGATTTTCTTTTGTCTAAATGCAAAACAGGCAATC 302
 Db |||||||
 546 ACTGTGATGTGTTGATCAGTCTGAAGATATTCCTGCACTCCCAATGAATAGACATTC 487
 QY 303 ATAGCCTTTGTGTTTAAAGCAAAAACATTTCTCTCCGATTCATCCCATTC 352
 Db |||||||
 486 AAAGCTTTGGAGTTTCCAGTCCCAATTCGTCCTCTCTTTAGTCCAGTC 437

```

RESULT 4
AAF22305_02
Continuation (3 of 11) of AAF22305 from base 200001 (Arabidopsis thaliana chromosome 4)
WP Sequence split into 11 fragments LOCUS AAF22305 Accession Aaf22305
WP Fragment Name Begin End
WP AAF22305_00 1 110000
WP AAF22305_01 100001 210000
WP AAF22305_02 200001 310000
WP AAF22305_03 300001 410000
WP AAF22305_04 400001 510000
WP AAF22305_05 500001 610000
WP AAF22305_06 600001 710000
WP AAF22305_07 700001 810000
WP AAF22305_08 800001 910000
WP AAF22305_09 900001 1010000
WP AAF22305_10 1000001 1082138

Query Match 2.5%; Score 54; DB 3; Length 110000;
Best Local Similarity 51.7%; Pred. No. 0.0018;
Matches 123; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 114 TCTCCGATGACACACATGTTTTCATTCGGAAGATTCGAAAGAGTGCATAAGAAATATG 173
DB 31784 TCCTCATGTTCTGTTCTTCATCTACATGATTTCAAACCTGAGATGCTAACATGCG 31843

QY 174 ATTTTGGATCTTTTCACTCGTTGGTGGCTTCATGAGTGAGCTCAAGAGTCTCTCAATA 233
DB 31844 ATTCTGGATCTTTTGACACTACTTGTCCCTTCATATGCTTTAGCAAGTTTATCCCATGCT 31903

QY 234 TCAAAAGCCGAATCACAATATGAATGTGATGATTTGATTTGTTCTAATGCACAAAC 293
DB 31904 TCITTAGCTGATTCACAGTTTGAATCGCTTGAATGATTTGATTTCAACGAAGTTGAT 31963

QY 294 AGGCAATCATAGCTTTGTTTAAAGCAAAACATTTCTCCGATTCATCCATT 351
DB 31964 ATCAAGACAGTGCCTTGAATTAGTTTGGCCCTTTGCTCTCTGATCATTCATT 32021

RESULT 5
AAF22305_03
Continuation (4 of 11) of AAF22305 from base 300001 (Arabidopsis thaliana chromosome 4)
WP Sequence split into 11 fragments LOCUS AAF22305 Accession Aaf22305
WP Fragment Name Begin End
WP AAF22305_00 1 110000
WP AAF22305_01 100001 210000
WP AAF22305_02 200001 310000
WP AAF22305_03 300001 410000
WP AAF22305_04 400001 510000
WP AAF22305_05 500001 610000
WP AAF22305_06 600001 710000
WP AAF22305_07 700001 810000
WP AAF22305_08 800001 910000
WP AAF22305_09 900001 1010000
WP AAF22305_10 1000001 1082138

Query Match 2.5%; Score 53.4; DB 3; Length 110000;
Best Local Similarity 50.2%; Pred. No. 0.0026;
Matches 132; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 107 GTACATGTCCTCGATGAGTACATGTTGTTTCATTTCGGAAGATTCGAAAGAGTGCATAAG 166
DB 14372 GTTGAAGTCATCCAGGATCTTTCAGCTTCATGTCAGATTCCTCAACTCTGAGCGAAG 14431

QY 167 AATATGATTTTGGATCTTTTCACTCGGTTTGGTGTGCTTCATGAGTGAGCTCAAGAGTCT 226
DB 14432 CATGTCCAGACAGTACGCTTCACATTTGTTGTCATTCATTCAGACTTCGAGAAATTC 14491

QY 227 CCAATATCAAAGCCGATCACAATGAATGATTTGATTTGATTTTGTCTAATGC 286
DB 14492 CCATACCTCTCTTGTGTAAGACATCCCTTGAACCTGAGTGAACTTATTCCTTAACAAAGA 14551

QY 287 ACAAAACAGGGCATTCATAGCCCTTTGTGTTTAAAGCAAAACATTTCTTCCGATTCATC 346

```

```

DB 14552 TCCAAAATCAGACAGTGCCTGAGAGTTGTGATTTGGCTGCTCTTCTCATCCGCAAT 14611

QY 347 CMTTCGCTCATCGGAGAGAAA 369
DB 14612 CCATTCAGTCTTGATTTGAAA 14634

RESULT 6
AAQ81475
ID AAQ81475 standard; DNA; 515 BP.
XX AAQ81475;
XX AC
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 04-SEP-1995 (first entry)
XX
XX HMG2 promoter III.
XX
XX HMG2; promoter; transgenic plant; post-harvest production;
KW 3-hydroxy-3-methylglutaryl-CoA-reductase; HMG2; wounding;
KW pathogen infection; pest infestation; inducible expression;
KW pollen-specific expression; disease-resistance; pesticide-resistance;
KW crop improvement; pDW101; ss.
XX
XX Lycopersicon esculentum; cv. VFNT Cherry.
XX
XX Key Location/Qualifiers
FT misc_feature 169..191
FT FT /*tag= a
FT FT /note= "bases at positions 169-191 form a 23-base
FT FT palindrome"
XX
XX W09503690-A1.
XX
XX 09-FEB-1995.
XX
XX 02-AUG-1994; 94WO-US008722.
XX
XX 02-AUG-1993; 93US-00100816.
XX 29-JUL-1994; 94US-00282581.
XX
XX (VIRG ) VIRGINIA TECH INTELLECTUAL PROPERTIES.
XX
XX Cramer CL, Weissenborn DL;
XX
XX WPI; 1995-081942/11.
XX
XX HMG2 promoter expression systems - used for inducible prodn. of gene
XX prods. in plants and plant cell cultures, partic. post-harvest prodn.
XX
XX Disclosure; Fig 6; 110pp; English.
XX
XX A 2.5 kb EcoRI fragment of the HMG2 gene (HMG2)-containing clone pTH295
XX was inserted into the EcoRI site of Bluescript SK-vector to generate
XX pDW101. The 515 bp sequence given in AAQ81475 is from the 5' end of the
XX 2.5 kb EcoRI insert of pDW101. The HMG2 promoter can be used to control
XX gene expression in transgenic plants or cell cultures in response to
XX wounding, elicitors, pest infestation, pathogen infection, etc. (Updated
XX on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct
XX PR field.) (Updated on 16-OCT-2003 to standardise OS field)
XX
XX Sequence 515 BP; 125 A; 121 C; 87 G; 182 T; 0 U; 0 Other;
XX
Query Match 2.4%; Score 52.6; DB 2; Length 515;
Best Local Similarity 51.5%; Pred. No. 0.00028;
Matches 121; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 135 TTCATTTCGGAAGATTCGAAAGAGTGCATAAGATTTGATTTGGATTTCTTCACTCGG 194
DB 122 TTCATTTCGGAAGTCTCATATCGTGAGTGAGCATCTTGGATTTCTTCACTTGT 181

```



```
XX AC F73552;
XX DT 20-NOV-2003 (first entry)
XX DE Staphylococcus aureus DNA #1232.
XX KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
XX KW enzymatic assay; antibiotic target; gene; ds.
XX OS Staphylococcus aureus.
XX PN WO200294868-A2.
XX PD 28-NOV-2002.
XX PF 27-MAR-2002; 2002WO-IB002637.
XX PR 27-MAR-2001; 2001GB-00007661.
XX PA (CHIR-) CHIRON SPA.
XX PI Masignani V, Mora M, Scarselli M;
XX WPI; 2003-120786/11.
XX DR P-PSDB; ABM71992.
XX ST New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX FS Claim 6; SEQ ID NO 2463; 49pp; English.
XX CC The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus genes of the invention
XX SQ Sequence 1722 BP; 657 A; 240 C; 310 G; 515 T; 0 U; 0 Other;
Query Match 1.9%; Score 41.2; DB 7; Length 1722;
Best Local Similarity 49.1%; Pred. No. 0.96;
Matches 109; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
QY 564 TGTAGGATCAGAGTGGCACTAAGAGAGAGAGGGGGAGTGAATTAGTGCAGTGAATAAAC 623
DB 1293 TGTTCCTTAGAAGTCCAAATAGGTACGGAAGTATTTGAAGTGGTGTGGTAATCAACG 1352
QY 624 TTATAGTTTAAATGATTCGTAATACGAGAGATTCGTTTTATAGTAACTTCGAG 683
DB 1353 TAATAAATATGATATATCTGCAGAAACAGTGCATTAATCATTTGAATAATGATTTTA 1412
QY 684 TAGATGAACAAACAAAGTTAAACAGTAGTGTAATAACAAATTCGGGAAAGTAAGACTCA 743
DB 1413 TAGATTACACGTTATTTCCAAAGAGACTTTTAATAACATTCGTTATATCTTACAGTAT 1472
QY 744 CACATTCAGGAACATACCAATTTAAAGTGGTTCGGTCAAAA 785
DB 1473 TACATTAACGATTCAATCTCATTCATGAGTTAAAGGTAAAAA 1514
RESULT 13
ID AA74565
XX AA74565 standard; DNA; 2881 BP.
XX AC AA74565;
XX
```

```
DT 16-MAR-1999 (first entry)
XX Staphylococcus aureus contig SEQ ID #254.
XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX OS Staphylococcus aureus.
XX Key Location/Qualifiers
FH misc_feature 541..600
FT /tag= a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They are
FT included to maintain the nucleotide numbering given in
FT the specification for this DNA sequence"
FT 2341..2400
FT misc_feature /tag= b
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They are
FT included to maintain the nucleotide numbering given in
FT the specification for this DNA sequence"
XX EP786519-A2.
XX 30-JUL-1997.
XX 07-JAN-1997; 97EP-00100117.
XX 05-JAN-1996; 96US-0009861P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
WPI; 1997-374922/35.
XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
PT stored on computer readable medium and used in the production of anti-
PT S.aureus vaccines.
XX Claim 1; Page 1047-1049; 3271pp; English.
XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or
CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the computer
XX readable medium
XX Sequence 2881 BP; 1049 A; 376 C; 505 G; 826 T; 0 U; 125 Other;
SQ
Query Match 1.9%; Score 41.2; DB 2; Length 2881;
Best Local Similarity 49.1%; Pred. No. 1.2;
Matches 109; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
QY 564 TGTAGGATCAGAGTGGCACTAAGAGAGAGGGGGAGTGAATTAGTGCAGTGAATAAAC 623
DB 1481 TGTTCCTTAGAAGTCCAAATAGGTACGGAAGTATTTGAAGTGGTGTGGTAATCAACG 1540
```


Mon Sep 20 12:58:17 2004

Db 5082 CAACAACTACTACTAAATTCGAAATTTACTAAATTTACTAAATTTAAATC 5033

Search completed: September 18, 2004, 19:20:43
Job time : 765.294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 11:51:07 ; Search time 152.359 Seconds
(without alignments)
7853.010 Million cell updates/sec

Title: US-09-892-635A-44

Perfect score: 2156

Sequence: 1 ggatcccaacttttagaat.....agcctcctcattcactatgg 2156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2.6/prodata/2/ina/5A COMB.seq.*
2: /cgn2.6/prodata/2/ina/5B COMB.seq.*
3: /cgn2.6/prodata/2/ina/6A COMB.seq.*
4: /cgn2.6/prodata/2/ina/6B COMB.seq.*
5: /cgn2.6/prodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2.6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	62.8	2.9	2417	4	US-09-254-776B-3
C 2	52.6	2.4	515	1	US-08-282-581-5
C 3	52.6	2.4	515	1	US-08-550-544-5
C 4	48.4	2.2	7218	1	US-08-232-463-14
C 5	45	2.1	7218	1	US-08-232-463-14
C 6	42.8	2.0	832	4	US-09-621-976-2813
C 7	41.2	1.9	2881	4	US-08-956-171E-254
C 8	37.6	1.7	5335	4	US-10-204-708-17
C 9	37.6	1.7	6317	4	US-10-204-708-11
C 10	36.6	1.7	399	4	US-09-621-976-8976
C 11	36.6	1.7	2871	4	US-09-023-655-768
C 12	36.2	1.7	1797	4	US-09-134-000C-1552
C 13	36.2	1.7	6113	4	US-10-204-708-13
C 14	35.8	1.7	3001	4	US-09-539-333D-192
C 15	35.8	1.7	6124	4	US-08-213-419B-3
C 16	35.8	1.7	21234	4	US-09-810-671-3
C 17	35.8	1.7	21234	4	US-10-109-854-3
C 18	35.6	1.7	1704	4	US-09-134-001C-925
C 19	35.6	1.7	168575	4	US-09-426-290-1
C 20	35.4	1.6	4185	4	US-09-417-485D-7
C 21	35.4	1.6	10640	4	US-09-417-485D-5
C 22	35.2	1.6	11050	4	US-10-204-708-86
C 23	35	1.6	6656	4	US-10-204-708-76
C 24	35	1.6	10467	4	US-10-204-708-2
C 25	35	1.6	72928	3	US-09-009-913-1
C 26	35	1.6	1664976	4	US-08-916-421B-1
C 27	34.8	1.6	3440	1	US-08-471-791-27

C 28 34.8 1.6 3440 5 PCT-US91-01746-27 Sequence 27, Appl
C 29 34.8 1.6 5562 4 US-10-204-708-64 Sequence 64, Appl
C 30 34.6 1.6 627 4 US-09-495-050A-74 Sequence 74, Appl
C 31 34.6 1.6 6020 4 US-10-204-708-8 Sequence 8, Appl
C 32 34.4 1.6 832 4 US-09-621-976-2813 Sequence 2813, Ap
C 33 34.4 1.6 887 3 US-09-602-848-3 Sequence 3, Appl
C 34 34.4 1.6 99500 4 US-09-798-096-10 Sequence 10, Appl
C 35 34.2 1.6 337 1 US-09-985-799-79 Sequence 79, Appl
C 36 34.2 1.6 337 1 US-08-594-031-79 Sequence 79, Appl
C 37 34.2 1.6 2427 4 US-09-134-001C-268 Sequence 268, App
C 38 34.2 1.6 4673 1 US-07-638-431-1 Sequence 1, Appl
C 39 34.2 1.6 4673 5 PCT-US92-00018-1 Sequence 1, Appl
C 40 34.2 1.6 5152 4 US-10-204-708-74 Sequence 74, Appl
C 41 34.2 1.6 580073 4 US-08-545-528D-1 Sequence 1, Appl
C 42 34.2 1.6 640681 4 US-09-790-988-1 Sequence 1, Appl
C 43 34 1.6 505 4 US-09-621-976-15639 Sequence 15639, A
C 44 34 1.6 576 4 US-09-621-976-16902 Sequence 16902, A
C 45 34 1.6 1497 4 US-09-220-132-94 Sequence 94, Appl

ALIGNMENTS

RESULT 1

US-09-254-776B-3/c

; Sequence 3, Application US/09254776B

; Patent No. 6559359

; GENERAL INFORMATION:

; APPLICANT: Laten, Howard

; TITLE OF INVENTION: PLANT RETROVIRAL POLYNUCLEOTIDES AND METHODS FOR USE THEREOF

; FILE REFERENCE: 27013/33479A

; CURRENT APPLICATION NUMBER: US/09/254, 776B

; CURRENT FILING DATE: 1999-03-09

; NUMBER OF SEQ ID NOS: 86

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3

; LENGTH: 2417

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: SIRE-1 cdna clone

US-09-254-776B-3

Query Match

Best Local Similarity 2.9%; Score 62.8; DB 4; Length 2417;

Matches 177; Conservative 0; Mismatches 172; Indels 1; Gaps 1;

QY 4 TCCCACTTTTAGGATGGATCTTAAATTTTAGTTATAGTTCAAGTTAGAAAATCT 63

Db 786 TCAATCTCTTAGGCAAGGATCTGAGGATCTTTCTCACCGCTTTTCATCAGTCATTCT 727

QY 64 TTACCAAGAGCTTTGAGTCCATGATGACATC- CGTGAACGGTGACATGTCCTCGATG 122

Db 726 TCTCCCAAGCAGTCGACGATTCGCAATTCGCAAGATTCATGTGAAGTCATGAATA 667

QY 123 GACTCAGCTGGTTTCATTCGGAAGAGTTCGAAAGAGTGCATAAGAAATATTTGGAT 182

Db 666 CACTCTTCCTCTCATCTTCAGATTTTCGAATTTTGTGGCCAAATAGTTGCAATCTGGAC 607

QY 183 TCTTTCACTCGGTTGGTGGCTTCATGAGTCAAGTCTCTCAATATCAAAAGCC 242

Db 606 ATCTTCACTTTGGAGGTTCTTCATGAGTGGTTTTCAGATCTCCCATCTCTTGGCC 547

QY 243 GAATCACAATTAATGATGATTAATTCATTTTGTCTTAATGCACAAACAGGCATTC 302

Db 546 ACTGTGCAATGTTTCATCAGTCGAGATATTCCTGCACTCCATTCGATTAAGCATTC 487

QY 303 ATAGCTTTGTTTAAAGCAAAAACATTTCTCCGATTCATCCATTC 352

Db 486 AAAGCTTTGGAGTTTCCAAGTCCCAATTCGTTCTTCTTTTAGTCCAGTC 437

COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REFERENCE/DOCKET NUMBER: 29,768
REGISTRATION NUMBER: 30472/114 IMMU
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 2.2%; Score 48.4; DB 1; Length 7218;
Best Local Similarity 12.0%; Pred. No. 0.00052;

Matches 61; Conservative 214; Mismatches 235; Indels 0; Gaps 0;

1454 TGTGGGAGTTTACGAGTAGCGAACCTTCTCGGTGATCTCGCAACCGCCGATGATC 1513
Db 979 TGGGCTACTATACTATTTTCTCTGTTGCGATACGCTCACAGAAATTAATTCGAGC 1038
1514 TCTTCGCGACACTTTCGAAACTTCGAAACTTCGAACTTCGATTTCTCGGTGTTCCGAC 1573
Db 1039 TTGGCTGCGAGTTCGAGGAGCTTGCGATTTTTTTTTTTTTTTTTTTTTTTTTT 1098
1574 AGCATCTCAAGAACTTCGAGTCTCGACTCTCGATGTCATCGAACTTCGACTCCGCTAGGCT 1633
Db 1099 YY 1158
1634 TGCTTTATATTTTCAGGCTATCATAGTTAATCTCATACATCTTAACCAATATGATTT 1593
Db 1159 YY 1218
1694 AGAATAATTAACCATCAATGATTTTCATCAAAATTCGACATTCGACATCAACAACTCCGT 1753
Db 1219 YY 1278
1754 ACTCAATAACCATCAGGCTATAGTACGATCTATCTAGTATCGGCTAGGCTAGGTT 1813
Db 1279 YY 1338
1814 AGCGAGTCATGATCCAGGCTGCTACTATTGCGGACACATGATCCCTTATCCAAATC 1873
Db 1339 YY 1398
1874 CAGTCTTCTCACTCTTACGCTACCGCTCTCTTTTATTACTTTTGAAGAATTCGA 1933
Db 1399 YY 1458
1934 AATCAAAACAGATACAAAATAACACGGTGA 1963
Db 1459 ACTACTTGATAGTAGTAAATACAGTGA 1488

RESULT 5
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOMLOPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Hardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 2.1%; Score 45; DB 1; Length 7218;
Best Local Similarity 7.0%; Pred. No. 0.0058;

Matches 30; Conservative 212; Mismatches 187; Indels 0; Gaps 0;

QY 320 AGCAAAACATCTTCTCGATTCATCCCATTCGCTCATCGGAGAGAAAATTTTGA 379
Db 1493 AGGCATCACTGAATTAATCTATCTGCAAGTAGTTAAAGAGATAGAAGATTTGCTACR 1434
QY 380 TCCATTTTCGACAATAGACCAAGCTCGAAATCCATGGAATCAGGAAGATCCTCATATG 439
Db 1433 RRR 1374
QY 440 AGTTTTCACATATGTAATTCGACTCATTAACATAGGTGATGTGTAATGAAATGACC 499
Db 1373 RRR 1314
QY 500 CTCATGCSCTATCTCTTGGGTATTAACCAAAATATGAGAGTACGCTTCTCTGATAC 559
Db 1313 RRR 1254
QY 560 CAATTCTTAGGATCAGAGTGGCAGTAAAGAGAGGGGGAGTGAATAGTGCAGTGAATTA 619
Db 560 CAATTCTTAGGATCAGAGTGGCAGTAAAGAGAGGGGGAGTGAATAGTGCAGTGAATTA 619

GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 17
; LENGTH: 5535
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-17

Query Match 1.7%; Score 37.6; DB 4; Length 5535;
Best Local Similarity 57.8%; Pred. No. 0.92;
Matches 67; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1615 AACTTGACTCGGTAGCTTGTATATTTTCAGGCTATCATAGTTAATCCCTACATCT 1674
Db 2300 AACTTACTTACTCAATACCTTTTCTTTCCCTCAATTAATAATCCAAATCT 2241

QY 1675 TAACTCAATAATGATAGATTAAATTAACCATCAATGATTCATCATCAAAA 1730
Db 2240 AAATAAATACGATACTCATACCTATAATCCCAACATTTAAATAACCAATAA 2185

RESULT 9
US-10-204-708-11
; Sequence 11, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 11
; LENGTH: 6317
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-11

Query Match 1.7%; Score 37; DB 4; Length 6317;
Best Local Similarity 48.8%; Pred. No. 1.5;
Matches 100; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 564 TGTTAGGATCAGAGTGGCCTAAGAGAGGGGGGAGTGAATTAGTCAGTCGATTAAAC 623
Db 2087 TGGTGGTTTATGTTTGTAAATTAGTAATTCGGGAGGTAGAGATTAGAGATTATTGAGT 2146

QY 624 TTATAAGTTTAAATAAATCGTAAATACGAGAAGATTTGTTTTTAATAGTAACCTTGAG 683
Db 2147 TTAAGAGTTTGGAGTTAGTTTGGTAAATATAGGAAGATTTATTTTATAAAAAATTAAA 2206

QY 684 TAGATGAACCAACCAAGTTAACAGTAGCTGTAATAACAAATTCGGGAAGTAAGAACTCA 743
Db 2207 AAAAAAATAATAGTAGGGTGTGTATGTTGTGTTTGTAGATATTATACGAGTT 2266

QY 744 CACATTCAAGGAACATACCAATTTA 768
Db 2267 GAGTAGGAGAATTATTTAGTTTA 2291

RESULT 10
US-09-621-976-8976
; Sequence 8976, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENEST.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8976
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 1.7%; Score 36.6; DB 4; Length 399;
Best Local Similarity 10.0%; Pred. No. 0.37;
Matches 30; Conservative 141; Mismatches 130; Indels 0; Gaps 0;

QY 575 GAGTGGCCTAAGAGAGGGGGGAGTGAATTAGTCAGTCGATTAAACCTTATAAGTTTA 634
Db 44 RRRRRRAWWKSMCMWKKSKSWRSWGMWTKEMKGRGAASWAGYMSWMTYTERWRYR 103

QY 635 AATAATCGAATTCGTAATACAGAGAAGATTTGTTTTTAATAGTAACCTTGAGTAGATAAAC 694
Db 104 YRKACTKWRAAGMWGKAGWAWAYAKWYMAWRRTAMKYWMMKSKSRMRERRAWYAWM 163

QY 695 CAAAAGTTAACAGTAGTGTAAATAACAAATTCGGGAAGTAAGAACTCACACATTCAAG 754
Db 164 YWMAERTWGRASCYRGAYMASAGYMYMMYMMRKKWMSAGWMSMKWTRRCASYS 223

QY 755 AACATACCAATTTAAAGTGGTTCCGTCAAAATGACCTACATCCACTTGTGAAGCCCTCTT 814
Db 224 CWSSYCMGAKMMYWKTSRWSYVYSSYRKYRSCCWSMSYWKTYRWSYSCASY 283

QY 815 CGAAGAGGCTCCCACTCCACTAGCAATACATCTTTGAAGGGGAGGACAAATACCTCTC 874
Db 284 SYKTRASCMMCCMKRKNMMYMKWTYCKTSAMRYRSCYSAKWRWRWMCAY 343

QY 875 T 875
Db 344 W 344

RESULT 11
US-09-023-655-768/c
; Sequence 768, Application US/09023655

```
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA: US/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 768:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTTUT01
; CLONE: 602872
; US-09-023-655-768

Query Match          1.7%; Score 36.6; DB 4; Length 2871;
Best Local Similarity 45.5%; Pred. No. 1.2;
Matches 126; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 510 ATCTCTCTGGTATTAAACCAATATGAGAGTGAGCCTTGCTCTGATACCAATGTTAG 569
Db 717 ATATTACAGGTCACAGAAACAAAGCTGTGCCAGTCATGTTGGACAATAACGTTTG 658

QY 570 GATCAGAGTGGCACTAAGAGAGGGGGAGTGAATAGTCAGTGGATTAAACCTTATAA 629
Db 657 GGGTCAGACGGGAAAGAGGAGAAAGAAAGAAAGAGAGAGAGAAATACTAATTT 598

QY 630 GTTTAAATGAATTCGTTAAATACGAGAAGATTCGTTTAAATAGTAATCTGAGTAGATG 689
Db 597 CTGGAACACACATTTGGCTTAACTGCCAAATAAAGGCTTTTCGGAGAAATGAAAGCCT 538

QY 690 AAAACCAAGTTAAGTAGTGTAAATTAACAAATTCGGGAAAGTAAGAACTCACACATT 749
Db 537 ATAATCAGATTAGTGTGCATTAACACACAGCTGACACCAACAGCAATATCTATCATAT 478

QY 750 CAAGGAACATACCAATTTAAAGTGTTCGGTCAAAAT 786
Db 477 TAATANTCTTCNATTATAGATCTATCTGTCACAAAT 441

RESULT 12
US-09-134-000C-1552

; Sequence 1552, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1552
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-1552

Query Match          1.7%; Score 36.2; DB 4; Length 1797;
Best Local Similarity 49.2%; Pred. No. 1.2;
Matches 95; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 922 AAGAAGGAGGTGAACATGCAAGCAATTGCAAAACAAAGACTTGCTCTAAAGACTTTGCTAAGGC 981
Db 761 AAGAAGGAGACCAAGAAGCTAATGAATTATGGAATGTTCCGTAGCGAGTCTCTAAAAG 820

QY 982 TTTTCTCTCAATCTATTGCTTCTCAAAAGTTGTTATCTCTGCTGAGAAATGAGGGTAT 1041
Db 821 AATTGTATAAAATCTATTCCATGTTAGAAGTTGAATTTGATCTTATAATGGCGAAGCCT 880

QY 1042 TTATAGACCCCAAGAGGATTTAAATTTGGGTCCTCAAAATTTGCAATGCTCTTTGGGTTCCCG 1101
Db 381 TTTATAATGATAAATGGAATGTAATTTGTTACTTTATTAGAGAAAACATTTATTACAG 940

QY 1102 AGTTTGGCGGTGC 1114
Db 941 AAAATCAAGGTGC 953

RESULT 13
US-10-204-708-13
; Sequence 13, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEX, Alexander
; APPLICANT: PIERPUBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 13
; LENGTH: 6113
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-204-708-13

Query Match          1.7%; Score 36.2; DB 4; Length 6113;
```


Db 1857 ATATTTTAGAATATTAAATTAATTCCTTAGAATAACTATATAATTTGATAATCCGAATCT 1916
Qy 718 AACAAATTCGGGAAAGTAAGAACTCACACATTCAGGAACATACCAATTTTAAAGT 772
Db 1917 TAAAATGTTACAAAAATGAGAAAGTAAAAAACTCCTTATTATATATATATATAAT 1971

Search completed: September 18, 2004, 14:23:48
Job time : 155.359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 14:07:43 ; Search time 967.103 Seconds
(without alignments)
11252.468 Million cell updates/sec

Title: US-09-892-635A-44

Perfect score: 2156

Sequence: 1 ggatccaaacttttagaat.....agcctctcattcattatgg 2156

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA: *

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2154.6	99.9	2156	13	US-09-892-635A-44
2	2139	99.2	2160	13	Sequence 44, Appl
3	2125.6	98.6	4924	13	Sequence 45, Appl
4	127.6	5.9	78333	16	Sequence 27, Appl
5	92	4.3	58985	16	Sequence 3, Appl
6	81.4	3.8	2760	13	Sequence 2, Appl
7	75	3.5	809	9	Sequence 108338,
8	73.4	3.4	5016	17	Sequence 48, Appl
9	73.4	3.4	5304	17	Sequence 10782, A
10	71.8	3.3	3843	17	Sequence 44132, A
11	71.8	3.3	4326	17	Sequence 10117, A
12	71.8	3.3	4428	17	Sequence 9179, Ap
13	71.8	3.3	4840	17	Sequence 44137, A
14	71.8	3.3	10369	17	Sequence 9223, Ap
					Sequence 10605, A

c 15	70.2	3.3	4002	17	US-10-437-963-9586	Sequence 9586, Ap
c 16	70.2	3.3	4428	17	US-10-437-963-9621	Sequence 9621, Ap
c 17	69.6	3.2	2160	17	US-10-437-963-101688	Sequence 101688, Ap
c 18	68.6	3.2	4884	17	US-10-437-963-9664	Sequence 9664, Ap
c 19	68	3.2	846	17	US-10-437-963-96678	Sequence 96678, A
c 20	67.4	3.1	4716	17	US-10-437-963-10750	Sequence 10750, A
c 21	67	3.1	4125	17	US-10-437-963-9312	Sequence 9312, Ap
c 22	67	3.1	4620	17	US-10-437-963-10128	Sequence 10128, A
c 23	67	3.1	5370	17	US-10-437-963-9697	Sequence 9697, Ap
c 24	67	3.1	6036	17	US-10-437-963-44135	Sequence 44135, A
c 25	65.4	3.0	5092	17	US-10-437-963-10162	Sequence 10162, A
c 26	64.8	3.0	1731	17	US-10-437-963-9134	Sequence 9134, Ap
c 27	62.8	2.9	2417	15	US-10-334-703-3	Sequence 3, Appli
c 28	62.8	2.9	2801	16	US-10-396-122-3	Sequence 3, Appli
c 29	62.8	2.9	2801	13	US-10-424-599-108650	Sequence 108650, Ap
c 30	62.8	2.9	9399	16	US-10-396-122-93	Sequence 93, Appl
c 31	62.6	2.9	1777	17	US-10-437-963-34059	Sequence 34059, A
c 32	61.2	2.8	5371	17	US-10-437-963-9498	Sequence 9498, Ap
c 33	61.2	2.8	9114	17	US-10-437-963-10728	Sequence 10728, A
c 34	61	2.8	4107	17	US-10-437-963-9190	Sequence 9190, Ap
c 35	60.2	2.8	3026	17	US-10-437-963-44142	Sequence 44142, A
c 36	59.6	2.8	9072	16	US-10-396-122-87	Sequence 87, Appl
c 37	59	2.7	3463	17	US-10-437-963-10660	Sequence 10660, A
c 38	58	2.7	5928	17	US-10-437-963-44133	Sequence 44133, A
c 39	57.6	2.7	4503	17	US-10-437-963-8782	Sequence 8782, Ap
c 40	56.6	2.6	7628	13	US-10-424-599-108640	Sequence 108640, Ap
c 41	56	2.6	3127	17	US-10-437-963-44134	Sequence 44134, A
c 42	53.4	2.5	4382	17	US-10-437-963-44138	Sequence 44138, A
c 43	53.4	2.5	5856	17	US-10-437-963-23904	Sequence 23904, A
c 44	53.2	2.5	13744	17	US-10-437-963-11092	Sequence 11092, A
c 45	52.8	2.4	5148	17	US-10-437-963-34840	Sequence 34840, A

ALIGNMENTS

RESULT 1

US-09-892-635A-44
; Sequence 44, Application US/09892635A
; Publication No. US20030226175A1
; GENERAL INFORMATION:
; APPLICANT: May, Gregory D.
; APPLICANT: Glendennen, Stephanie K.
; APPLICANT: Mason, Hugh S.
; APPLICANT: Lim, Miguel A. Gomez
; APPLICANT: Arntzen, Charles J.
; TITLE OF INVENTION: DNA Regulatory Elements Associated with Fruit Development
; FILE REFERENCE: 031998-007
; CURRENT APPLICATION NUMBER: US/09/892, 635A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/160,351
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: US 60/060,062
; PRIOR FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Musa acuminata
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 879
; OTHER INFORMATION: n = A,T,C or G
US-09-892-635A-44

Query Match 99.9%; Score 2154.6; DB 13; Length 2156;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCCAACTTTTAGGAATGGATCTTAAATTTTGTATAGTTCAAAGTTAGAAAAA 60
|||||
Db 1 GGATCCCAACTTTTAGGAATGGATCTTAAATTTTGTATAGTTCAAAGTTAGAAAAA 60
|||||

QY 61 TCCTTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAAACGGTGTACATGCTCCGA 120
Db 61 TCCTTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAAACGGTGTACATGCTCCGA 120
QY 121 TGGACTCAGCTGGTGTTCATTCGGAAAAAGTTCCGAAAGAGTGATAGAAATATTGATTTGG 180
Db 121 TGGACTCAGCTGGTGTTCATTCGGAAAAAGTTCCGAAAGAGTGATAGAAATATTGATTTGG 180
QY 181 ATTCCTTTCACTCGGTGGTCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAG 240
Db 181 ATTCCTTTCACTCGGTGGTCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAG 240
QY 241 CCGAATCAAAATTTGAATGTGATGAAATTCATTTTGTCTAAATGCGACAAAACAGGGCAT 300
Db 241 CCGAATCAAAATTTGAATGTGATGAAATTCATTTTGTCTAAATGCGACAAAACAGGGCAT 300
QY 301 TCATAGCCTTTGTGTTTAAAGCAAAACAAATTCCTCCGATTCATFCCCATTCGGTCAATCG 360
Db 301 TCATAGCCTTTGTGTTTAAAGCAAAACAAATTCCTCCGATTCATFCCCATTCGGTCAATCG 360
QY 361 GAAGAGAAAATTTTGAATTCATTTTCGA CAATAGACCAAGCTCGAAAATCCCATGGAAA 420
Db 361 GAAGAGAAAATTTTGAATTCATTTTCGA CAATAGACCAAGCTCGAAAATCCCATGGAAA 420
QY 421 TGAGAGAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACTTAACACATAGGTG 480
Db 421 TGAGAGAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACTTAACACATAGGTG 480
QY 481 GATGTGTAATGAATGACCCCTCATGCSCTATCTCTCTTGGGTATTAACCAAAATATGAGA 540
Db 481 GATGTGTAATGAATGACCCCTCATGCSCTATCTCTCTTGGGTATTAACCAAAATATGAGA 540
QY 541 GTGAGCCTTGCTCTGATACCAATTTGTTAGGATCAGAGTGGCACTAAGAGAGGGGGAGT 600
Db 541 GTGAGCCTTGCTCTGATACCAATTTGTTAGGATCAGAGTGGCACTAAGAGAGGGGGAGT 600
QY 601 GAAATAGTGCAGTGAATTAACACTTATAAGTTTAAAAATGAATTCGTAANTACGAGAGA 660
Db 601 GAAATAGTGCAGTGAATTAACACTTATAAGTTTAAAAATGAATTCGTAANTACGAGAGA 660
QY 661 TTTCTGTTTAAATAGTAACTTGAAGTAACTCAAAAGTTAACAGTAGTGTAAATTAAC 720
Db 661 TTTCTGTTTAAATAGTAACTTGAAGTAACTCAAAAGTTAACAGTAGTGTAAATTAAC 720
QY 721 AATTTCCGGAAGTAAGAACTCACATTCGAGGAAACATACCAATTTAAAGTGGTTCGGT 780
Db 721 AATTTCCGGAAGTAAGAACTCACATTCGAGGAAACATACCAATTTAAAGTGGTTCGGT 780
QY 781 CAABAATGACCTACATCCACTTGTGAAGCTTCTTCGAGAGGCTCCCAACTTCCACTAGC 840
Db 781 CAABAATGACCTACATCCACTTGTGAAGCTTCTTCGAGAGGCTCCCAACTTCCACTAGC 840
QY 841 AAATCACTTTGAAGGGGAAGCAAAATACCTCTCTTACNACCTTTTACAAATGGTTCATAC 900
Db 841 AAATCACTTTGAAGGGGAAGCAAAATACCTCTCTTACNACCTTTTACAAATGGTTCATAC 900
QY 901 TCTTACAAATTTCAACGAGAAAGAGAGGTGAACATGCAAGCAATTTGAAAACAAAGACT 960
Db 901 TCTTACAAATTTCAACGAGAAAGAGAGGTGAACATGCAAGCAATTTGAAAACAAAGACT 960
QY 961 TGCTAAAGACATTTGCTAAGGCTTTTCTCAATCTATTGCTTCTCAAAAGTTGATTTCT 1020
Db 961 TGCTAAAGACATTTGCTAAGGCTTTTCTCAATCTATTGCTTCTCAAAAGTTGATTTCT 1020
QY 1021 CTGCTGAGAAATGAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGCTCCAAAT 1080
Db 1021 CTGCTGAGAAATGAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGCTCCAAAT 1080
QY 1081 TCGAATGCTCTTGGGTTCCCGAGGTTGCGGTGCCACCGCTGTGACGTGTTGACACTGG 1140
Db 1081 TCGAATGCTCTTGGGTTCCCGAGGTTGCGGTGCCACCGCTGTGACGTGTTGACACTGG 1140

RESULT 2

US-09-892-635A-45

; Sequence 45, Application US/09892635A

; Publication No. US20030226175A1

QY 1141 ACAGTGTAAGCGGTGCCACCGCGGACCTCTCGGGTGTGGCGGTGCGACCGCCTAG 1200
Db 1141 ACAGTGTAAGCGGTGCCACCGCGGACCTCTCGGGTGTGGCGGTGCGACCGCCTAG 1200
QY 1201 ACITTTTTCAGTCTCAGTGGTTGGATTCCAAACCTTGACCCAAACCAAGTCCGAACTCGGTTCC 1260
Db 1201 ACITTTTTCAGTCTCAGTGGTTGGATTCCAAACCTTGACCCAAACCAAGTCCGAACTCGGTTCC 1260
QY 1261 AATTGACCGGTAAACCGGATTAAGGATTAACCTTAAATCCTTAACCTTAATATATGCAAA 1320
Db 1261 AATTGACCGGTAAACCGGATTAAGGATTAACCTTAAATCCTTAACCTTAATATATGCAAA 1320
QY 1321 CTACGCAACTGAAATATAGTCTTAAGCAAGTTTTTAAACCGGCAAAACGTCGAGTCTCTT 1380
Db 1321 CTACGCAACTGAAATATAGTCTTAAGCAAGTTTTTAAACCGGCAAAACGTCGAGTCTCTT 1380
QY 1381 CCGCGCATCTTTCCGGCAGACTTCTGATATATACCTTTGGATTTCTTCTAGCGGACTCCTAGT 1440
Db 1381 CCGCGCATCTTTCCGGCAGACTTCTGATATATACCTTTGGATTTCTTCTAGCGGACTCCTAGT 1440
QY 1441 AGGGTCCGATCTTTGTGCGGAGTTTAGGAGTAGCGGAACCTTCTCGGTGATCTCCGCAA 1500
Db 1441 AGGGTCCGATCTTTGTGCGGAGTTTAGGAGTAGCGGAACCTTCTCGGTGATCTCCGCAA 1500
QY 1501 ACCGCCGATGATCTCTTCCGCGAGACTTTTCGAAAACTTCGACAAAGTCCCGGATTTCTTCTC 1560
Db 1501 ACCGCCGATGATCTCTTCCGCGAGACTTTTCGAAAACTTCGACAAAGTCCCGGATTTCTTCTC 1560
QY 1561 GGTTGGTTCGACAGCATCTCTAAACGAACTTCGGAATTCCTTGAATGTCCATCGAATCTG 1620
Db 1561 GGTTGGTTCGACAGCATCTCTAAACGAACTTCGGAATTCCTTGAATGTCCATCGAATCTG 1620
QY 1621 ACTCCGTTAGGCTTGTCTTATATTTTTCAGGATCATAGTTTAACTTACATTAACATC 1680
Db 1621 ACTCCGTTAGGCTTGTCTTATATTTTTCAGGATCATAGTTTAACTTACATTAACATC 1680
QY 1681 AATAATATGGATTAAGATTAATTAACCCATCAATTAATTAATTAATTAATTAATTAATTAAT 1740
Db 1681 AATAATATGGATTAAGATTAATTAACCCATCAATTAATTAATTAATTAATTAATTAATTAAT 1740
QY 1741 AACAAACATCCGTACTCAATTAACCCATCAAGCTATAGTTAGTGAATCTTACTGTGATC 1800
Db 1741 AACAAACATCCGTACTCAATTAACCCATCAAGCTATAGTTAGTGAATCTTACTGTGATC 1800
QY 1801 CGTACGTGAAGTTAGCGAGTCAATCCAGTCCGTCGTCACCTTATTTGCGCGACAGGATC 1860
Db 1801 CGTACGTGAAGTTAGCGAGTCAATCCAGTCCGTCGTCACCTTATTTGCGCGACAGGATC 1860
QY 1861 CTTTATCCAAATCCAGTCTTCTCAACTCTTCTAGCCTACCGGTCTCTTTTTTATTAATCTT 1920
Db 1861 CTTTATCCAAATCCAGTCTTCTCAACTCTTCTAGCCTACCGGTCTCTTTTTTATTAATCTT 1920
QY 1921 TTGAAAGAAATTCAAATCAAAACAGATCAAAATTAACCGGTGAGACATGTGACATGCTA 1980
Db 1921 TTGAAAGAAATTCAAATCAAAACAGATCAAAATTAACCGGTGAGACATGTGACATGCTA 1980
QY 1981 GTCTCTGGAAGAGCAATTAATTCGCGCATCCAGAGGTCGTGAGTTCATCAACCCACTTTT 2040
Db 1981 GTCTCTGGAAGAGCAATTAATTCGCGCATCCAGAGGTCGTGAGTTCATCAACCCACTTTT 2040
QY 2041 TCCTACATACCATGTCGATCGCTTTTGTGTATGACAGACCAACCAAGCTTGCCTTTGGT 2100
Db 2041 TCCTACATACCATGTCGATCGCTTTTGTGTATGACAGACCAACCAAGCTTGCCTTTGGT 2100
QY 2101 TGTGCTTAACAG 2156
Db 2101 TGTGCTTAACAG 2156

; GENERAL INFORMATION:

; APPLICANT: May, Gregory D.
; APPLICANT: Clendennen, Stephanie K.
; APPLICANT: Mason, Hugh S.
; APPLICANT: Lim, Miguel A. Gomez
; APPLICANT: Aritzzen, Charles J.
; TITLE OF INVENTION: DNA Regulatory Elements Associated with Fruit Development
; FILE REFERENCE: 031998-007
; CURRENT APPLICATION NUMBER: US/09/892,635A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/160,351
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: US 60/060,062
; PRIOR FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Musa acuminata
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 883
; OTHER INFORMATION: n = A,T,C or G
; -09-892-635A-45

Query Match 99.2%; Score 2139; DB 13; Length 2160;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2155; Conservative 0; Mismatches 1; Indels 4; Gaps 1;

QY	1	GGATCCCAACTTTTAGGAATGGATCTTAAATTTTATAGTTTCAAAAGTTAGAAAAA	60
DB	1	GGATCCCAACTTTTAGGAATGGATCTTAAATTTTATAGTTTCAAAAGTTAGAAAAA	60
QY	61	TCATTACCAAGAGCTTTGAGTCCATTGATGACATCCCGTGAACGGTGTACATGCTCCGA	120
DB	61	TCATTACCAAGAGCTTTGAGTCCATTGATGACATCCCGTGAACGGTGTACATGCTCCGA	120
QY	121	TGGACTCAGTGGTTTCATTCGGAAGAGTTTGAAGAGTGCATAGAGTATGATTTGG	180
DB	121	TGGACTCAGTGGTTTCATTCGGAAGAGTTTGAAGAGTGCATAGAGTATGATTTGG	180
QY	181	ATTCTTTCACTCGGTTGGTGCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAG	240
DB	181	ATTCTTTCACTCGGTTGGTGCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAG	240
QY	241	CCGATCACAATTTGAATGTGATTTGATTTGCTTAATGCAACAAACAGGGCAT	300
DB	241	CCGATCACAATTTGAATGTGATTTGATTTGCTTAATGCAACAAACAGGGCAT	300
QY	301	TCATAGCCTTTGTTTAAAGCAAAACATTTCTCCGATTCATCCCATTCGCTCATCG	360
DB	301	TCATAGCCTTTGTTTAAAGCAAAACATTTCTCCGATTCATCCCATTCGCTCATCG	360
QY	361	GAAGAGAAATTTTGAATCCATTTTCGACATAGACCAAGCTCGAAATC---CATG	416
DB	361	GAAGAGAAATTTTGAATCCATTTTCGACATAGACCAAGCTCGAAATC---CATG	420
QY	417	GAATGAGGAAGATTCCTCATATAGATTTTCAATATACATGTAATTCGACTCATTAACATA	476
DB	421	GAATGAGGAAGATTCCTCATATAGATTTTCAATATACATGTAATTCGACTCATTAACATA	480
QY	477	GGTGGATGTGATGAATGACCTCATGCGCTATCTCTCTGGGTATTAACCAATAT	536
DB	481	GGTGGATGTGATGAATGACCTCATGCGCTATCTCTCTGGGTATTAACCAATAT	540
QY	537	GAGAGTGAGCCTTGCTCTGATACCAATTTGTAGGATCAGAGTGGCCTTAAGAGAGGGGG	596
DB	541	GAGAGTGAGCCTTGCTCTGATACCAATTTGTAGGATCAGAGTGGCCTTAAGAGAGGGGG	600
QY	597	GAGTGAATAGTCAGTGGATTTAAACCTTATAAGTTTAAAGTTTAAAGTTTAAAGTTACGAG	656
DB	601	GAGTGAATAGTCAGTGGATTTAAACCTTATAAGTTTAAAGTTTAAAGTTTAAAGTTACGAG	660

QY	657	AAGATTTTCGTTTAAATAGTAACTTAGTAGATGAAACCAAAAGTTTAAACAGTAGTAA	716
DB	661	AAGATTTTCGTTTAAATAGTAACTTAGTAGATGAAACCAAAAGTTTAAACAGTAGTAA	720
QY	717	TAAACAATTTTCGGAAGAGTAAAGACTCACAATTCAGAGAAACATACCAATTTAAAGTGGTT	776
DB	721	TAAACAATTTTCGGAAGAGTAAAGACTCACAATTCAGAGAAACATACCAATTTAAAGTGGTT	780
QY	777	CGGTCAAAATGACCTACATCCACTTGTGAAGCCCTTCTTGAAGAGGCTCCCAACTCCAC	836
DB	781	CGGTCAAAATGACCTACATCCACTTGTGAAGCCCTTCTTGAAGAGGCTCCCAACTCCAC	840
QY	837	TAGCAATTCACCTTTGAAGGGAAGGACAAATACCTCTCTACACCTTTTACAATGGTTC	896
DB	841	TAGCAATTCACCTTTGAAGGGAAGGACAAATACCTCTCTACACCTTTTACAATGGTTC	900
QY	897	ATACTCTTACAAATTTTCAACGAGAAAGAGGAGTGAACATGCAAGCAATTTGAAACAA	956
DB	901	ATACTCTTACAAATTTTCAACGAGAAAGAGGAGTGAACATGCAAGCAATTTGAAACAA	960
QY	957	GACTTGCTAAAGACTTTTGTGAAGCTTTTTTCTCAATCTATTGCTTCTCAAAAGTTGTA	1016
DB	961	GACTTGCTAAAGACTTTTGTGAAGCTTTTTTCTCAATCTATTGCTTCTCAAAAGTTGTA	1020
QY	1017	TTCTCTGCTGAGATTTGAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGCTCCA	1076
DB	1021	TTCTCTGCTGAGATTTGAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGCTCCA	1080
QY	1077	AAATTTGCAATGCTTTGGGTTCCCGAGTTGCGGGTGCACCGCTCTGAGTGTTCACA	1136
DB	1081	AAATTTGCAATGCTTTGGGTTCCCGAGTTGCGGGTGCACCGCTCTGAGTGTTCACA	1140
QY	1137	CTGACAGAGTGTATAGCGGTGCCACCGCCGACCTCTCGGGTGTGGCGGTGCGCACCGC	1196
DB	1141	CTGACAGAGTGTATAGCGGTGCCACCGCCGACCTCTCGGGTGTGGCGGTGCGCACCGC	1200
QY	1197	CTAGACTTTTTCAGCTCAGTGGTTGATTTCAAACTTGACCCAAACAGTCCGACTCGG	1256
DB	1201	CTAGACTTTTTCAGCTCAGTGGTTGATTTCAAACTTGACCCAAACAGTCCGACTCGG	1260
QY	1257	GTCCAAATGACCGGTAAACCGGATTTAGAGTAAACCTTAATCTTAACCCCTAATATATG	1316
DB	1261	GTCCAAATGACCGGTAAACCGGATTTAGAGTAAACCTTAATCTTAACCCCTAATATATG	1320
QY	1317	CAAACTACGCACTGAAATATAGTCTTAAGCAAGTTTTTAAACCGGCAACGTCGAGTCT	1376
DB	1321	CAAACTACGCACTGAAATATAGTCTTAAGCAAGTTTTTAAACCGGCAACGTCGAGTCT	1380
QY	1377	TTCTCCGCGATCTTTTCGCGAGACTTCTGATATACCTTTGGATTTCTTCTAGCGGACTCC	1436
DB	1381	TTCTCCGCGATCTTTTCGCGAGACTTCTGATATACCTTTGGATTTCTTCTAGCGGACTCC	1440
QY	1437	TAGTAGGTTCCCGATCTTGTGGCGAGTTTAGCGAGTAGCGGAACCTTCTCGGTGATCTCC	1496
DB	1441	TAGTAGGTTCCCGATCTTGTGGCGAGTTTAGCGAGTAGCGGAACCTTCTCGGTGATCTCC	1500
QY	1497	GCAACCGCGGATGATCTCTTCGCGAGACTTTTCAAACTTGACAAAGTCCCGGATTTCT	1556
DB	1501	GCAACCGCGGATGATCTCTTCGCGAGACTTTTCAAACTTGACAAAGTCCCGGATTTCT	1560
QY	1557	TTCTCCGTTGGTCCGACAGACTCTTAAACGAACTTCGACTCTTGAATGTCATCGAA	1616
DB	1561	TTCTCCGTTGGTCCGACAGACTCTTAAACGAACTTCGACTCTTGAATGTCATCGAA	1620
QY	1617	CTTGACTCCGTTAGGCTTCTTTATATTTTTCAGGCTATCATAGTAACTCTTACATCTTA	1676
DB	1621	CTTGACTCCGTTAGGCTTCTTTATATTTTTCAGGCTATCATAGTAACTCTTACATCTTA	1680
QY	1677	ACTCAATAATAGGATTTAGATTTAAATTAACCCCAATGATTTTCAATCATCAAAATTCGAC	1736
DB	1681	ACTCAATAATAGGATTTAGATTTAAATTAACCCCAATGATTTTCAATCATCAAAATTCGAC	1740

```
QY 1737 ATTCAACAAACATCGTACTCAATAAACCATCAGGCTATAGTTAGCTGACTATCTACTGT 1736
|
|
|
Db 1741 ATTCAACAAACATCCGTACTCAATAAACCCATCAGGCTATAGTTAGCTGACTATCTACTGT 1800
|
|
|
QY 1797 GATCCGTACGTTAGAGTTAGCGAGTCAATGATCCAGTCTGTGTCACTTATTTGGCCGAACACG 1856
|
|
|
Db 1801 GATCCGTACGTTAGAGTTAGCGAGTCAATGATCCAGTCTGTGTCACTTATTTGGCCGAACACG 1860
|
|
|
QY 1857 TATCCCTTATCCAAATCCAGTCTCTCAACTCTTCTAGCTTACCGCTCTCTTTTATTTAT 1916
|
|
|
Db 1861 TATCCCTTATCCAAATCCAGTCTCTCAACTCTTCTAGCTTACCGCTCTCTTTTATTTAT 1920
|
|
|
QY 1917 ACTTTTGAAGAATTCAAAATCAAAACAGATACAAATAAACAACGCTGAGACACTGTGACAT 1976
|
|
|
Db 1921 ACTTTTGAAGAATTCAAAATCAAAACAGATACAAATAAACAACGCTGAGACACTGTGACAT 1980
|
|
|
QY 1977 GCTAGTCTCTGGAAGCAATTAAATTCGCGCATCCACAGACGTCGTCAAGCTTCAACACCCAC 2036
|
|
|
Db 1981 GCTAGTCTCTGGAAGCAATTAAATTCGCGCATCCACAGACGTCGTCAAGCTTCAACACCCAC 2040
|
|
|
QY 2037 TTTTTCCTACATACATGTCGATGCTTGTGATGACAGACACACCAACGCTTGCCTT 2096
|
|
|
Db 2041 TTTTTCCTACATACATGTCGATGCTTGTGATGACAGACACACCAACGCTTGCCTT 2100
|
|
|
QY 2097 TGGTTGTCCCTAAACAGAGAGAGAGAGACAGACACGATAGCTTCTCATTTCACTATGG 2156
|
|
|
Db 2101 TGGTTGTCCCTAAACAGAGAGAGAGAGACAGACACGATAGCTTCTCATTTCACTATGG 2160
|
|
|
RESULT 3
US-09-892-635a-27
; Sequence 27, Application US/09892635A
; Publication No. US20030226175A1
; GENERAL INFORMATION:
; APPLICANT: May, Gregory D.
; APPLICANT: Clendenen, Stephanie K.
; APPLICANT: Mason, Hugh S.
; APPLICANT: Lim, Miguel A. Gomez
; APPLICANT: Arntzen, Charles J.
; TITLE OF INVENTION: DNA Regulatory Elements Associated with Fruit Development
; FILE REFERENCE: 031998-007
; CURRENT APPLICATION NUMBER: US/09/892,635A
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/160,351
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: US 60/060,062
; PRIOR FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 4924
; TYPE: DNA
; ORGANISM: Musa acuminata
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 879..3691., 4119
; OTHER INFORMATION: n = A,T,C or G
US-09-892-635a-27
Query Match 98.6%; Score 2125.6; DB 13; Length 4924;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2152; Conservative 0; Mismatches 0; Indels 5; Gaps 2;
QY 1 GGATCCCAACTTTTAGGAATGGATCTTAAATTTAGTTATTAAGTTCAAAGTTAGAAAA 60
Db 1 GGATCCCAACTTTTAGGAATGGATCTTAAATTTAGTTATTAAGTTCAAAGTTAGAAAA 60
QY 61 TCTTTTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGACGCTGATGCTCCGA 120
Db 61 TCTTTTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGACGCTGATGCTCCGA 120
QY 121 TGGACTCACTTGGTTTCATTCGAAAAGTTCCGAAAGAGTGATCAATAAGATTTTGG 180
|
|
|
```

```
Db 121 TGGACTCACTTGGTTTCATTCGAAAAGTTCCGAAAGAGTGATCAATAAGATTTTGG 180
QY 181 ATTCTTTCACCTCGGTTGGTGCCTTCATGAGTGACCTCAAGAGTCTCTCCAAATATCAAAAG 240
|
|
|
Db 181 ATTCTTTCACCTCGGTTGGTGCCTTCATGAGTGACCTCAAGAGTCTCTCCAAATATCAAAAG 240
|
|
|
QY 241 CCGAATCACAAATTTGAAATGTGATTGAATTCATTTTGTCTAAATGTCACAAAACAGGGCAT 300
|
|
|
Db 241 CCGAATCACAAATTTGAAATGTGATTGAATTCATTTTGTCTAAATGTCACAAAACAGGGCAT 300
|
|
|
QY 301 TCATAGCCTTTGTCTTTAAAGCAAAAAACATTCCTTCGATTCATCCCATTCGTCATCG 360
|
|
|
Db 301 TCATAGCCTTTGTCTTTAAAGCAAAAAACATTCCTTCGATTCATCCCATTCGTCATCG 360
|
|
|
QY 361 GAAGAGAAAAATTTTGAATTCATATGAGTTTCCAAATACATGTAATTCGACTCATTTAAACATAGGTG 480
|
|
|
Db 361 GAAGAGAAAAATTTTGAATTCATATGAGTTTCCAAATACATGTAATTCGACTCATTTAAACATAGGTG 480
|
|
|
QY 421 TGAGGAAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACTCATTTAAACATAGGTG 480
|
|
|
Db 421 TGAGGAAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACTCATTTAAACATAGGTG 480
|
|
|
QY 481 GATGTGTAATGAAATGACCCCTCATGCSCTATCTCTCTTGGGTATTTAAACCAAAATATGAGA 540
|
|
|
Db 481 GATGTGTAATGAAATGACCCCTCATGCSCTATCTCTCTTGGGTATTTAAACCAAAATATGAGA 540
|
|
|
QY 541 GTGAGCCTTGTCTGATACCAATTTGTTAGGATCAGAGTGGCCTAAAGAGAGGGGGGAGT 600
|
|
|
Db 541 GTGAGCCTTGTCTGATACCAATTTGTTAGGATCAGAGTGGCCTAAAGAGAGGGGGGAGT 600
|
|
|
QY 601 GAATTAGTGACGTGGATTAAACCTTTATAAGTTTAAAAATGAATTCGTTAAATACGAGAAGA 660
|
|
|
Db 601 GAATTAGTGACGTGGATTAAACCTTTATAAGTTTAAAAATGAATTCGTTAAATACGAGAAGA 660
|
|
|
QY 661 TTTCCGTTTAAATAGTAACCTTCAGTAGATGAAAAACAAAAGTTAACAGTAGTGTAAATAAC 720
|
|
|
Db 661 TTTCCGTTTAAATAGTAACCTTCAGTAGATGAAAAACAAAAGTTAACAGTAGTGTAAATAAC 720
|
|
|
QY 721 AATTTCCGGAAAGTTAAGAACTCACACATTCAAGGAACATACCAATTTAAAGTGTTCGGT 780
|
|
|
Db 721 AATTTCCGGAAAGTTAAGAACTCACACATTCAAGGAACATACCAATTTAAAGTGTTCGGT 780
|
|
|
QY 781 CAAAATGACCTACATCCACTTTGTGAAGCCTTCTTCGAAGAGGCTCCCAACTTCCACTAGC 840
|
|
|
Db 781 CAAAATGACCTACATCCACTTTGTGAAGCCTTCTTCGAAGAGGCTCCCAACTTCCACTAGC 840
|
|
|
QY 841 AAATCACTTTGAAGGGGGAAGACAAATACCTCTCTTACNAGCTTTTACAAATGGTTCATAC 900
|
|
|
Db 841 AAATCACTTTGAAGGGGGAAGACAAATACCTCTCTTACNAGCTTTTACAAATGGTTCATAC 900
|
|
|
QY 901 TCTTACAAATTTTCAACGAGAAAGAGGAGTGAACTGCAAGCAATTTGAAACCAAGACT 960
|
|
|
Db 901 TCTTACAAATTTTCAACGAGAAAGAGGAGTGAACTGCAAGCAATTTGAAACCAAGACT 960
|
|
|
QY 961 TGCTAAAGACTTTGCTAAGGCTTTTTTTTCTCAATCTATTGTCTCTCAAAAGTTGATTTCT 1020
|
|
|
Db 961 TGCTAAAGACTTTGCTAAGGCTTTTTTTTCTCAATCTATTGTCTCTCAAAAGTTGATTTCT 1020
|
|
|
QY 1021 CTGCTGAGAAATGAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGCTCCAAAT 1080
|
|
|
Db 1021 CTGCTGAGAAATGAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGCTCCAAAT 1080
|
|
|
QY 1081 TCGAATGCTCTTGGGTTCCCGAGGTTCCCGGTGCCCGCTGTCAAGTGTTCACACTGG 1140
|
|
|
Db 1081 TCGAATGCTCTTGGGTTCCCGAGGTTCCCGGTGCCCGCTGTCAAGTGTTCACACTGG 1140
|
|
|
QY 1141 ACAGTGTACTAGCGGTGCCCGCGGACCTCTCGGGTGTGGGGGTGTCACCCGCTAG 1200
|
|
|
Db 1141 ACAGTGTACTAGCGGTGCCCGCGGACCTCTCGGGTGTGGGGGTGTCACCCGCTAG 1200
|
|
|
QY 1201 ACTTTTTCAGCTCACTGTTGGATTCCAAACTTGACCCCAACCACTGCGAATCCGGTCC 1260
|
|
|
Db 1201 ACTTTTTCAGCTCACTGTTGGATTCCAAACTTGACCCCAACCACTGCGAATCCGGTCC 1260
|
|
|
```



```
; NAME/KEY: misc feature
; LOCATION: (2145)..(2320)
; OTHER INFORMATION: n denotes unsequenced nucleotides
US-10-298-122-2

Query Match      4.3%; Score 92; DB 16; Length 58965;
Best Local Similarity 57.8%; Pred. No. 1.2e-12;
Matches 186; Conservative 0; Mismatches 130; Indels 6; Gaps 1;

Qy 181 ATTCTTTCACTCGGTGGTGGCTTCATGAGTGACCTCAAGAGTCTCTCCAAATATCAAAAG 240
Db 38730 ATTCTCTCTCTTTGGTAGTCTCTCATACACAAATTTTGAGTTTCTCCAAATTTCAATAG 38789

Qy 241 CCGAATCACAATGAATGATGAATTCATTTTGTCTAATGCACAAAACAGGGCAT 300
Db 38790 CATTAGAACAAATTTGAAACTCTATGAAATTCCTTTTATCTAAGGCACAAAATAAGCAT 38849

Qy 301 TCATAGCCTTTGTGTTTAAAGCAAAACATTCCTTCGGATTCATCCCATTCGGTCAATCG 360
Db 38950 TCATAGCTTTGGAATTTAAACACATATTTTCTTATCTAATCACTTAAGCATTCITTC 38909

Qy 361 GAAGAGAAAATTTTGAATCCATTTTCGACAAATAGACCAAGCTCGAAATCCCATGGAAA 420
Db 38910 TAGGA-----ATGAGTAAATCATCACAAACAACTTCCCAAATTTCAATATCTAAAGCTT 38963

Qy 421 TGAGGAAGATCCTCATATGAGTTTTCATATCATGTAATTCGACTCATTAACATAGTGG 480
Db 38964 GTAAATAAATTCATCTAGTTTTCCTCAATAAGGATAATCATTTCCATCTAAATAATGGAG 39023

Qy 481 GATGTGTAATGAATGACCCCTC 502
Db 39024 GTCITGTGATAGATTGCTCTTC 39045

RESULT 6
US-10-424-599-108338
; Sequence 108338, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 108338
; LENGTH: 2760
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_68844C.1
US-10-424-599-108338

Query Match      3.8%; Score 81.4; DB 13; Length 2760;
Best Local Similarity 53.8%; Pred. No. 1.2e-10;
Matches 189; Conservative 0; Mismatches 161; Indels 1; Gaps 1;

Qy 1  GGATCCCACTTTTAGGATGATCTTAAATTTTAGTTATAGTTCAAGTTGAGTAAAGAAA 60
Db 1829 GGTTCGCCATCTCTACTTAAACATCTTAACCTTGTGTTATAGATCTCATATTTGAAAT 1889

Qy 61 TCTTTACCAAGAGCTTTGAGTCCATGATGACATCCGTAAGCGGTGACATGTCCTCG- 119
Db 1889 TCTTTGCCCTAAGGCTGCTAGATGATTTACTATATGTAATCTCTTTTGCATACCTCTGA 1948

Qy 120 ATGGACTCACTTGGTTTCATTCGGAAGCTTCGAAGAGTGCATAGATATTTGATTG 179
Db 1949 ATATTTTTCATTGTCATTCATCTTAAATAATTCATCTCATGAGTTAGTGCATTTATCCTA 2008

180 GATTCCTTTCACTCGGTGGTGGCTTCATGAGTGACCTCAAGAGTCTCTCCAAATATCAAAA 239
2009 GAICTTTTAACTTTGATTTCTCTTCATGTTGTTAATCGAAGAGTGTCCACATTTCTTTA 2068
240 GCGGAATCACAATTTGAAATGTGATGAATTCATTTTGTCTAATGCACAAAACAGGGCA 299
2069 GCATCTTTCAATTTGAAACCCCTGAAATATTCATCCATTCCTAGGCGAGATGTTATTATG 2128
300 TTCATAGCCTTTGTGTTTAAAGCAAAACATTCCTCTCCGATTCATCCCAT 350
2129 TTTTGGCTTTTAAAGTTGATTTGATCTTTTCTATCTCTCTCTCTCAGACCAT 2179

RESULT 7
US-09-966-881-48
; Sequence 48, Application US/09966881
; Patent No. US20020120960A1
; GENERAL INFORMATION:
; APPLICANT: Seymour, Graham
; Bird, Colin
; Medina-Suarez, Rosybel
; TITLE OF INVENTION: Genetic control Of Fruit Ripening
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zeneca Ag Products Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/966,881
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,860
; FILING DATE: 29-Mar-1999
; APPLICATION NUMBER: GB 9618862.8
; FILING DATE: 10-SEP-1996
; APPLICATION NUMBER: GB 9708366.1
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: PCT/GB97/02424
; FILING DATE: 08-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: SEE 50183/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-1699
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 809 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: U-D86
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-09-966-881-48

Query Match      3.5%; Score 75; DB 9; Length 809;
Best Local Similarity 100.0%; Pred. No. 2.9e-09;
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2082 CCACAAAGCTTGCTTTGGTTGCTTAAACAGAGAGAGAGACACAGACCGATAGCCT 2141
Db 9 CCACAAAGCTTGCTTTGGTTGCTTAAACAGAGAGAGAGACACAGACCGATAGCCT 68
```


; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 44137

; LENGTH: 4428

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_47227C.1

US-10-437-963-44137

Query Match 3.3%; Score 71.8; DB 17; Length 4428;

Best Local Similarity 49.1%; Pred. No. 6.4e-08;

Matches 250; Conservative 1; Mismatches 248; Indels 10; Gaps 2;

QY 4 TCCCACTTTTAGGAATGGATCTTAAATTTTATAGTTTCAAGTTCAAGTTAGAAAATCT 63

Db 500 TCATACCTCTCTGATAGGCTCTAAGCATCTTTTGGCAACCTCAAGATCGGTGATTT 441

QY 64 TTACCAAGAGCTTTGAGTCATTCGATGACATCCGTGAAC-GGTGTACATGTCCTCGATG 122

Db 440 GCCCAAGCCCTTGAGATCAATTCACATGACATTAAGCTCCCATACATGTCATTCACA 381

QY 123 GACTCACTTGGTTTCATTCGAAAAGTTTCAAGAGTGCATAAGATATTTGTTGGAT 182

Db 380 CTCCTATGAGGCAACATGGAGATGTCATATTTGATCTTGAGAAAATGAAGCTTGGCA 321

QY 183 TCTTCACTCGGTGGTGGCTTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAAGCC 242

Db 320 TCCTTGTACTCACTTGTACCTCATGATCTCGGCAACTTTGTCAAATCTCATATGG 261

QY 243 GAATCAAAATGAAATGATGATTAATTCATTTTGTCTTAATGCACAAACAGGGCATTC 302

Db 260 GTCTCAAGGTGTCTCACTCTATCGAACTCTCTTGGCTCAAAGAGTTGAACAAGGCATTC 201

QY 303 ATAGCCTTTGTGTTTAAAGCAAAACATTTCTTCCGATTCATCCCATCGCTCATCGGA 362

Db 200 ATGCG-----TTGGGCAATGAGTTGGAGTTGGGTGATCAATCTCCGTCAGGGT 150

QY 363 AGAGAAAATTTTGAATCCATTTTCGACATAGACAAAGCTCGAAATCCATCGAAATG 422

Db 149 GTCCCGGTGATAGCAAAAGCTACATCCACAATCTCCAAATATAGAAGCTCATAGCTTTG 90

QY 423 AGAAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACTCATTAACATAGTGA 482

Db 89 AGGTGAGTAGACATTTTAAATTTTCCAAAGTGGAGTAGTTTGTGCCATTTAAACATGGGAGCC 30

QY 483 TGTGTAATGAAATGACCTCATGCSCTAT 511

Db 29 TTCCTACATGGTTCACTCGTTCGACAT 1

RESULT 13

US-10-437-963-9223

; Sequence 9223, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 9223

; LENGTH: 4840

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_1565C.1

US-10-437-963-9223

Query Match 3.3%; Score 71.8; DB 17; Length 4840;

Best Local Similarity 49.1%; Pred. No. 6.7e-08;

Matches 250; Conservative 1; Mismatches 248; Indels 10; Gaps 2;

QY 4 TCCCACTTTTAGGAATGGATCTTAAATTTTATAGTTTCAAGTTCAAGTTAGAAAATCT 63

Db 4341 TCATACCTCTCTCGGTAGGCTCTAAGCATCTTTTGGCAACCTCAAGATCGGTGATTT 4400

QY 64 TTACCAAGAGCTTTGAGTCATTCGATGACATCCGTGAAC-GGTGTACATGTCCTCGATG 122

Db 4401 GCCCAAGCCCTTGAGATCAATTCACATGACATTAAGCTCCCATACATGTCATTCACA 4460

QY 123 GACTCACTTGGTTTCATTCGAAAAGTTTCAAGAGTGCATAAGATATTTGTTGGAT 182

Db 4461 CTCCTATGAGGCAACATGGAGATGTCATATTTGATCTTGAGAAAATGAAGCTTGGTA 4520

QY 183 TCTTCACTCGGTGGTGGCTTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAAGCC 242

Db 4521 TCCTTGTACTCACTTGTACCTCATGATCTCGGCAACTTTGTCAAATCTCATATGG 4580

QY 243 GAATCAAAATGAAATGATGATTAATTCATTTTGTCTTAATGCACAAACAGGGCATTC 302

Db 4581 GTCTCAAGGTGTCTCACTCTATTTGAACCTCTTGGCTCAAAGAGTTGAACAAGGCATTC 4640

QY 303 ATAGCCTTTGTGTTTAAAGCAAAACATTTCTTCCGATTCATCCCATCGCTCATCGGA 362

Db 4641 ATGGCTTGGCAATGAGTTGAA-----GGTTGGGTGATCAATCTCCGTCAGGGT 4691

QY 363 AGAGAAAATTTTGAATCCATTTTCGACATAGACAAAGCTCGAAATCCATCGAAATG 422

Db 4692 GTTCGGTGATAGCAAAAGCTACATCCACAATCTCCAAATATGGAAGCTCATAGCTTT 4751

QY 423 AGAAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACTCATTAACATAGTGA 482

Db 4752 AGGTGAGTAGACATTTTAAATTTTCCAAAGTGGAGTAGTTTGTGCCATTTAAACATGGGAGCC 4811

QY 483 TGTGTAATGAAATGACCTCATGCSCTAT 511

Db 4812 TTCCTACATGGTTCACTCGTTCGACAT 4840

RESULT 14

US-10-437-963-10605

; Sequence 10605, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 10605

; LENGTH: 10369

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_1690C.1
US-10-437-963-10605

Query Match 3.3%; Score 71.8; DB 17; Length 10369;
Best Local Similarity 49.1%; Pred. No. 1.1e-07;
Matches 250; Conservative 1; Mismatches 248; Indels 10; Gaps 2;

Qy 4 TCCCAACTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAAAAGTTAGAAAATCT 63
Db 9870 TCATACITTCGGTAGGGCTTAAGCATCTTTGGGCAACCTCAAGATCGGTAGTTT 9929

Qy 64 TTACAAGAGCTTTAGTCCATTCATGACATCCCGTGAAC-GGTGTACATCTCTCCGATG 122
Db 9930 GCCCAAGCCCTTGAGATCATTTACAATGACATTTAGCCCTCCCATACATGTCATCACA 9989

Qy 123 GACTCACTTGTTCATTCGAAAAGTTGAAAGAGTGATGATGATGATGATGATGATGAT 182
Db 9990 CTCATGAGGCAACATGGAAATGTCTCATATTGGATCTTTGAGGAAATGAAGCTTGCA 10049

Qy 183 TCTTTCACTCGGTGGTGGCTTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAAGCC 242
Db 10050 TCTTTGACTCAGTGTACCTCATGATCTCCGCCAACTTGTCCAAATCTCATATGCG 10109

Qy 243 GAATCACAAAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
Db 10110 GTCTCAAGTTGCTCACTCTATCGAACTCCCTTGGCTCAAAAGAGTTGAACAAGGCATTC 10169

Qy 303 ATAGCCTTTGGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATTCGCTCATCGGA 362
Db 10170 ATGGC-----TTGGGCATTCAGTTGGAGTTGGCGGTATCAATCTCCGTCAGGGT 10220

Qy 363 AGAAAAATTTTGAATTCATTTTCGACATAGACCAAAAGCTCGAAATCCATGGAATG 422
Db 10221 GTGCGGTGATGGCAAGCTTACATCCAAATATCCAAATATGGAAGCTCATAGCTTTG 10280

Qy 423 AGGAAGATCCATATGATGATTTTCCAATACATGATGATGATGATGATGATGATGATG 482
Db 10281 AGGTGAGTAGACATTTAAATTTTCCAAGTGGAGTAGTTTGTGCCATTAAACATGGAGCC 10340

Qy 483 TGTGTAAATGAATGACCTCATGCTAT 511
Db 10341 TTCCCTACATGGTTACCTCGTTCGACAT 10369

RESULT 15

US-10-437-963-9586/c
Sequence 9586, Application US/10437963
Publication No. US2004012343A1

GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 9586
LENGTH: 4002
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_1598C.1
US-10-437-963-9586

Query Match 3.3%; Score 70.2; DB 17; Length 4002;
Best Local Similarity 48.9%; Pred. No. 1.6e-07;

Matches 249; Conservative 1; Mismatches 249; Indels 10; Gaps 2;

Qy 4 TCCCAACTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAAAAGTTAGAAAATCT 63
Db 500 TCATACITTCGGTAGGGCTCTAAGCATCTTTGGGCAACCTCAAGATCGGTAGTTT 441

Qy 64 TTACAAGAGCTTTAGTCCATTCATGACATCCCGTGAAC-GGTGTACATCTCTCCGATG 122
Db 440 GCCCAAGCCCTTGAGATCATTTACAATGACATTTAGCCCTCCCATACATGTCATCACA 381

Qy 123 GACTCACTTGTTCATTCGAAAAGTTGAAAGAGTGATGATGATGATGATGATGATGAT 182
Db 380 CTCATGAGGCAACATGGAAATGTCTCATATTGGATCTTTGAGGAAATGAAGCTTGCA 321

Qy 183 TCTTTCACTCGGTGGTGGCTTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAAGCC 242
Db 320 TCTTTGACTCACTGTGACCTCATGGATCTCCGCCAACTTGTTCCAAATCTCATATGCG 261

Qy 243 GAATCACAAAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 302
Db 260 GTCTCAAGTTGCTCACTCTATCGAACTCCCTTGGCTCAAGAGTTGAACAAGGCATTC 201

Qy 303 ATAGCCTTTGGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATTCGCTCATCGGA 362
Db 200 ATGGC-----TTGGGCATTCAGTTGGAGTTGTGGTATCAATCTCCGTCAGGGT 150

Qy 363 AGAAAAATTTTGAATTCATTTTCGACATAGACCAAAAGCTCGAAATCCATGGAATG 422
Db 149 GTGCGGTGATGGCAAGCTTACATCCAAATATCCAAATATGGAAGCTCATAGCTTTG 90

Qy 423 AGGAAGATCCATATGATGATTTTCCAATACATGATGATGATGATGATGATGATGATG 482
Db 89 AGGTGAGTAGACATTTAAATTTTCCAAGTGGAGTAGTTTGTGCCATTAAACATGGAGCC 30

Qy 483 TGTGTAAATGAATGACCTCATGCTAT 511
Db 29 TTCCCTACATGGTTACCTCGTTCGACAT 1

Search completed: September 18, 2004, 18:55:07
Job time : 971.103 secs

Fax: 814 865 9131
Email: cwd3@psu.edu or jhl10@psu.edu
The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (<http://pqm.cornell.edu>)

Project can be obtained at <http://esp.dic.ford.com>

RESULT 2	BH021865	790 bp	DNA	linear	GSS 29-JUN-2001
LOCUS	BH021865				
DEFINITION	GH MB0001K07f Gossypium hirsutum L. Gossypium hirsutum genomic clone GH_MB0001K07f, genomic survey sequence.				
ACCESSION	BH021865				
VERSION	BH021865.1	GI:14576153			
KEYWORDS	GSS.				
SOURCE	Gossypium hirsutum	(upland cotton)			
ORGANISM	Gossypium hirsutum				

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium. 1 (bases 1 to 790).
Tomkins, J.P., Peterson, D.G., Yang, T.J., Main, D., Wilkins, T.A., Paterson, A.H. and Wing, R.A.
Development of Genomic Resources for Cotton (*Gossypium hirsutum* L.): BAC Library Construction, Preliminary STC Analysis, and Identification of Clones Associated With Fiber Development

Fax: 604 656 4133
 Email: twing@clemson.edu
 Seq primer: TAATACGACTCTATAGGG
 Class: BAC ends
 High quality sequence stop: 532.
 Location/Qualifiers
 1. .790
 /organism="Gossypium hirsutum"
 /mol_type="genomic DNA"
 /cultivar="Maxxa"
 /db_xref="taxon:3635"
 /clone="GH_MBB000K07"
 /tissue_type="young leaves"
 /lab_host="E. coli"
 /clone_lib="Gossypium hirsutum L."
 /note="Vector: pCUGIBAC-1; Site 1: HindIII; Site 2: NotI;
 For more details on library preparation, ordering clones
 and sequence analysis see
 http://www.genome.clemson.edu/projects/stc/cotton/GH_MBB

RESULT 3					
CG823740/c					
LOCUS	CG823740	977 bp	DNA	linear	GSS 18-NOV-2003
DEFINITION	SOYDB38TH	LargeInsertSoybeanGenLib	Glycine max genomic clone		

H07C21:MTP9G4, genomic survey sequence.
 CG823740
 CG823740.1 GI:38284664
 GSS.
 Glycine max (soybean)
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 977)
 Shultz, J., Meksem, K., Shetty, J., Town, C.D., Koo, H., Potter, J.,
 Wakefield, K., Zhang, H., Wu, C. and Lightfoot, D.A.
 End sequencing of BACs comprising a provisional minimal tiling path
 from a fingerprint physical map of soybean (Glycine max) cultivar
 Forrest
 Unpublished (2003)
 Other GSSs: SOYDB38TV
 Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
 The Center of Excellence in Soybean Research, Teaching and
 Outreach, Southern Illinois University at Carbondale and Plant
 Genomics, The Institute for Genomic Research
 Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
 USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 618 453 1797
 Email: ga4082@siu.edu, jshultz@siu.edu, cdtown@tigr.org (URL:
 http://bioinformatics.siu.edu)
 Clones approximating a minimum tiling path were re-arrayed from the
 library master plates prior to sequencing.
 For purposes of clone identification each clone name is a
 concatenation of the original clone location and its new location
 in the re-arrayed sequencing plates.
 Seq primer: AATTAACCCCTCATTAAGG
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..977
 /organism="Glycine max"
 /mol_type="genomic DNA"
 /cultivar="Forrest"
 /db_xref="taxon:3847"
 /clone="H07C21:MTP9G4"
 /notes="Organ: Leaves: Vector: pCLD04541 (pBELOBAC11 EcoRI
 clones): Site: 1: BstXI; Soybean (Glycine max (L.) Merr.)
 cv. Forrest seeds were grown in greenhouse for fourteen
 days. Nuclei were isolated and embedded in agarose,
 restriction digested with Hind III BamHI or EcoRI, large
 size DNA fragments were ligated in vector V41 (pCLD04541)
 and electro transformed in DH10a cells. About 90,000
 clones from BAC libraries were fingerprinted with HindIII
 and Hae III. Version 2 (automatic build) Contigs were
 built from 78,001 fingerprints. Contigs were manually
 examined to find the best non redundant tile path through
 the contigs representing about 13,000 clones. The clones
 were end sequenced."

ORIGIN

Query Match 4.3%; Score 93.4; DB 29; Length 977;
 Best Local Similarity 51.7%; Pred. No. 6.4e-13;
 Matches 283; Conservative 0; Mismatches 236; Indels 28; Gaps 2;
 QY 1 GGATCCCACTTTTAGGAATGGATCTTAATAATTTTACTTATAGTCAAGTTAGAAAA 60
 Db
 572 GGTTGCCATCTCTACTTAAACATCTTAATACCTTTATATGATATATCTTCATTGGAAA 513
 QY 61 TCTTTACCAAGAGTTTGGTCCATTCATGACATCCCGTGAACGGTGTA-CATGTCCTCCG 119
 Db
 512 ATTTCCTTAAGATGCAAGATGATTATTATATGTTGTAACCTCTTTCATGCTCTGT 453
 QY 120 ATGGACTCACTTGGTTTCATTCGGAAAGTTCGAAAGAGTGCATAAGAAATATTGTTG 179
 Db
 452 ATACTTTCAATTGAATTCATTCTAAATAATTCATCTATGAGTGAATGATTTATCTTA 393

QY 180 GATTCTTTCACTCGGTGGTGCCTTCATGAGTGACCTCAAGAGTCTCTCCAAATATCAAAA 239
 Db
 392 GATATTTTCACTCTGTGTGGCCCTCATGTGTACTTGTAGGGTATCCCAATTTCTTTT 333
 QY 240 GCCGAATCACAATAATGAAATGTGATTAATTCATTTTGTCTAATGCACAAAAACAGGCA 299
 Db
 332 GCATTTTACAATTTGATACCCCTAAAGTACTCATCCATTCTTAAGGAGATGTAATTATA 273
 QY 300 TTCAATAGCCTTTGTGTTAAAGCAAAAAACATTTCTTCCGATTATCCCATTCGCTCATC 359
 Db
 272 TTTTGGCTTTTAAGTTGTATTGCACTAATGTCTTTCTCTTCTCACTCTATTCTCCCTA 213
 QY 360 GG-----AAGAGAAAAATTTTGAATCCATTTCCACA 392
 Db
 212 GGCTTTTCTAAGGTTTACTTCTACACCATATAGTGGGAATATAAGGCCCTACTTCTATT 153
 QY 393 ATAGACCAAGCTCGAAATCCATGGAATGAGGAAGATCCTCATATAGTTTTCCAAATAC 452
 Db
 152 GCATCCAGATATTAAATTTATTGCTCTCTATAAAATTTGCATATGGGTTTTCCAGTAA 93
 QY 453 ATGTAATTCGACTCATTAAACATAGGTGGATGTGTAATGAATGACCCCTATGCGCTATC 512
 Db
 92 TGGTAACCCACTCCATTGAAGATAGAGGCTAATTTATTGAATTTCTTCAGGAATAAAA 33
 QY 513 TCTCTTG 519
 Db 32 TGGTTTG 26

RESULT 4

LOCUS CC412052 964 bp DNA linear GSS 19-MAY-2003
 DEFINITION PUBJ02TB_ZM_0.6-1.0_KB_Zea_mays_genomic_clone_ZMMBta412P16,
 genomic survey sequence.

ACCESSION

VERSION CC412052

KEYWORDS

CC412052.1 GI:30892142

SOURCE

GSS.

ORGANISM

Zea mays

REFERENCE

1 (bases 1 to 964)

AUTHORS

Whitelaw, C.A., Fraser, C.M., Yuan, Y., Van Aken, S., Utterback, T.,
 Resnick, A., and San Miguel, P., Ma, J. and
 Bennetzen, J.

TITLE

Maize Genomics Consortium

JOURNAL

Unpublished (2003)

COMMENT

Other GSSs: PUBJ02TD

CONTACT

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel:

301-838-5843

Fax:

301-838-0208

Email:

whitelaw@tigr.org

Seq primer:

TR

Class:

sheared ends.

Location/Qualifiers

1..964

FEATURES

source

Query Match

4.3%; Score 93; DB 28; Length 964;

Best Local Similarity

53.6%; Pred. No. 8.1e-13;

Matches

262; Conservative 0; Mismatches 220; Indels 7; Gaps 3;

ORIGIN

Cot selected genomic DNA library"

QY 4 TCCCAACTTTTGGAGTGAATCTTAAATTTTGTAGTTTAACTTCAAGTTAGAAAATCT 63
 |||
 Db TCATATGCTTTGGAGAGATCTAAGGAATCTGTGAGTGAATCCACATCGATACATTA 150
 |||
 QY 64 TTACAAGAGCTTTGAGTCCATGATGACATCCGTGAACGG-TGTACATCTCTCCGATG 122
 |||
 Db 151 AAACCAAGTCCCTTGAGCTCATCTCATCTCATTTAAACCGGTTAAACATGTCGACACA 210
 |||
 QY 123 GACTCACTTGTTCATTCGAAAGTTCGAAAGAGTGCATAGAATATTCATTTGGAT 182
 |||
 Db 211 CTTTCATCCTCTTCATGAGAAATGCTCAAACTTTCCTTTGCGGTACATACAACTTAGCA 270
 |||
 QY 183 TCTTTCACTCGTTGGTGGCTTCATGAGTGAACCTCAAGAGTCCCTCCAAATATCAAAAGCC 242
 |||
 Db 271 CTCCTTCAAAATGTAGTTCCTTCATGAATTCCTAGCTTTTCCAAATCTCATGCGG 330
 |||
 QY 243 GAATCACAAATGAAATGTGATGAATTCATTTTGTCTAATGACACAAACAGGCGATTC 302
 |||
 Db 331 GTTGTGAGATTTTGTGCGGTGCAACTCAATTTATGTTCTAAAGCATCATAGATACATTC 390
 |||
 QY 303 ATAGCCTTGTGTTTAAAGCAAAACATCTCTCCGATTCATCCCATTCGCTCATCGGA 362
 |||
 Db 391 ATAG-CTTGATCATGTAGTAAATTTTCATATCAATTTGGTGAATCTTCATCTCTTT 449
 |||
 QY 363 AGAGAAATTTTGAATTCATTTTCGACATAGACCAAGCTCGAAATCCATCGAAATG 422
 |||
 Db 450 TAAATACAA-----CAATCCATCTTTACATAGAGCCAAATCTTTTCCCATAGCTTTA 504
 |||
 QY 423 AGAAGATCCCTCATATGATGTTTCCATATACATGAATTCGACTCATTTAAACATAGTGGA 482
 |||
 Db 505 AGATGCAATGACGTCTTATTTTCCAAATGGTGCATAAATTTTCCCATCAAAATGTTGTGCG 564
 |||
 QY 483 TGTGTAATG 491
 |||
 Db 565 TTCCCAATG 573

RESULT 5

CG822306/c
 LOCUS
 DEFINITION SOYEA21TV LargeInsertSoybeanGenLib Glycine max genomic clone
 H45P23:MP15C17, genomic survey sequence.

ACCESSION

CG822306

VERSION

CG822306.1 GI:38281855

KEYWORDS

GSS

SOURCE

Glycine max (soybean)

ORGANISM

REFERENCE

1 (bases 1 to 776)
 Shultz,J., Meksem,K., Shetty,J., Town,C.D., Koo,H., Potter,J.,
 Wakefield,K., Zhang,H., Wu,C. and Lightfoot,D.A.
 End sequencing of BACs comprising a provisional minimal tiling path
 from a fingerprint physical map of soybean (Glycine max) cultivar
 Forrest

JOURNAL

COMMENT

Unpublished (2003)
 Other GSSs: SOYEA21TH
 Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
 The Center of Excellence in Soybean Research, Teaching and
 Outreach, Southern Illinois University at Carbondale and Plant
 Genomics, The Institute for Genomic Research
 Room 176, Ag Building, Mail Code 4415, Carbondale, IL 62901-4415,
 USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 618 453 1797
 Email: ga4082@siu.edu, jleshultz@siu.edu, cdtown@igr.org (URL:
 http://bioinformatics.siu.edu)
 Clones approximating a minimum tiling path were re-arrayed from the
 library master plates prior to sequencing.
 For purposes of clone identification each clone name is a
 concatenation of the original clone location and its new location
 in the re-arrayed sequencing plates.

Seq primer: GTAATACGACTCACTATATAGGC
 Class: BAC ends
 FEATURES
 source
 1..776
 Location/Qualifiers
 /organism="Glycine max"
 /mol_type="genomic DNA"
 /cultivar="Forrest"
 /db_xref="taxon:3847"
 /clone="H45P23:MP15C17"
 /clone_lib="LargeInsertSoybeanGenLib"
 /note="Organ: Leaves; Vector: pCD04541 (pBELOBAC11 Ecori
 clones); Site 1: BstXI; Soybean (Glycine max (L.) Merr.)
 cv. Forrest seeds were grown in greenhouse for fourteen
 days. Nuclei were isolated and embedded in agarose.
 restriction digested with Hind III BamHI or EcoRI, large
 size DNA fragments were ligated in vector v41 (pCD04541)
 and electro transformed in DH10a cells. About 90,000
 clones from BAC libraries were fingerprinted with HindIII
 and Hae III. Version 2 (automatic build) Contigs were
 built from 78,001 fingerprints. Contigs were manually
 examined to find the best non redundant tile path through
 the contigs representing about 13,000 clones. The clones
 were end sequenced."

ORIGIN

Query Match 4.3%; Score 92.2; DB 29; Length 776;
 Best Local Similarity 51.8%; Pred. No. 1.3e-12;
 Matches 277; Conservative 0; Mismatches 233; Indels 25; Gaps 2;
 QY 1 GGATCCCAACCTTTAGGAATGGATCTTAAATTTTGTAGTTTAACTTCAAGTTAGAAA 60
 Db 568 GGTGGCCATCTCTACTTAAACATCTTAAACATCTTATATGAGATCTCTGTTTGA 509
 |||
 QY 61 TCATTACAGAGCTTTGAGTCCATTCATGATGACATCCGTGAAC--GGGTACATGTC 119
 |||
 Db 508 GTTCTTCTTAATGATGCAAGATGATTAACATATATGTGTGAATCTTTTCTGATATCT 449
 |||
 QY 120 ATGGAATCACTTGTGTTTTCATTCGCAAAAGTTTCGAAAGAGTGCATAAAGATATTC 179
 |||
 Db 448 ATACTTCATTTGTCTTCATCCCTAAATATTCATCTACTCATGATTAAGTATTTATC 389
 |||
 QY 180 GATCTTTCACTCGGTTGGTGGCTTCATGATGACCTCAAGAGTCCCTCCAAATATCA 239
 |||
 Db 388 GATCTTTGACATCAGTTGTTCCCTCATGTGTAACCTGTAGAGTGTCCACATATCC 329
 |||
 QY 240 GCCGAATCACAATTTGAATGTGATTCATTTTGTCTCTAATGACACAAACAGGCG 299
 |||
 Db 328 GCATCTTACAAATTTGAACCCCTAAATATTCATCCATTCCTAGGGCAGAAATATA 269
 |||
 QY 300 TTCATAGCTTTGTGTTTAAAGCAAAACATTTCTTCCGATTCATCCCATTCGCTCAT 359
 |||
 Db 268 TTTTATGCTTTAAATATATCTGACATTAATCTTTTCATCTTCAGACCATCTTCT 209
 |||
 QY 360 GG-----AAGAGAAAATTTTGAAATTCCTATGATTTTCCATACATG 455
 |||
 Db 148 TCCCAAAATGTTTAAAGTCTATTGCTCCTCAATGAAAATTTGCAITCGGGTTTCC 89
 |||
 QY 456 TAAATTCGACTCATTTAAACATAGTGGATGTCTAATGAATGAATGACCTCATGSC 510
 |||
 Db 88 TAACTCTCCATTAATAAATAGGAGGCTATTATTGAATTTCTCTCTGGAATA 34
 |||
 RESULT 6
 CG818965
 LOCUS
 DEFINITION SOYED06TV LargeInsertSoybeanGenLib Glycine max genomic clone
 H5G03:MP15B12, genomic survey sequence.
 ACCESSION
 CG818965
 VERSION
 CG818965.1 GI:38275242

CG818965 832 bp DNA linear GSS 18-NOV-2003
 SOYED06TV LargeInsertSoybeanGenLib Glycine max genomic clone
 H5G03:MP15B12, genomic survey sequence.

ACCESSION

CG818965

VERSION

CG818965.1

Matches 196; Conservative 0; Mismatches 156; Indels 1; Gaps 1;

QY 1 GGATCCCAACTTTAGGAATGATCTTAAATTTTAGTTTAAAGTTCAAGTTAGAAA 60
|||
Db 571 GGTGGCAATCTCTATTAAAGACATCTTAAAGACTTTATTATTAGATCTTCAATTTGAAA 512
|||
QY 61 TCTTTACCAAGAGCTTTGAGTCCATTGATGATCATCCGTTGAACCGTGTA-CATGTCCTCG 119
|||
Db 511 GTTTTCTAGAGATGCAAGATGATTAACTATGTGTGTAACCTCTTTTGATGATCTTGT 452
|||
QY 120 ATGACTCACTGTTTCATTCGGAAGTTGGAAGAGTGCAFAAGAAATATGATTTTG 179
|||
Db 451 ATACTTTTATTACATTCATCTCTAAAAGTTTCAATTCACGAGTTAAAGTCTTTATCCTA 392
|||
QY 180 GATCTTTCACTCGTGGTGGCTTCATGAGTGACCTCAAGAGTCTCCCAATATCAAAA 239
|||
Db 391 GATCTTTTAACTCTGTTGTGCTTCATGTTGTTACTTGTATGTATGCCACATATCCCTA 332
|||
QY 240 GCCGAATCACAAATTTGAATGTGATTTGAATTCATTTTGTCTAATGCACAAAACAGGGCA 299
|||
Db 331 GCACCTTTTACAATTTGAAACCTTAAAGTATTCACTATTCTTAAGGCAGATGTAATAA 272
|||
QY 300 TTAATAGCTTTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATC 352
|||
Db 271 TTTTGGCCTTTAAATTTATATTGTTACTAATCTTCTTCTCTCAGTCCATC 219
|||

RESULT 8
CC804388 787 bp DNA linear GSS 14-JUL-2003
LOCUS ih28all.g1 WGS-ZmaysU (DH10b or JM107MA2) Zea mays genomic clone
DEFINITION ih28all, genomic survey sequence.
ACCESSION CC804388
VERSION CC804388.1 GI:32629811
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 787)
Rabinowicz,P.D., O'Shaughnessy,A.L., Palmer,L., Balija,V.,
Dedhia,N., Katzenburger,F., King,L., Miller,B., Muller,S.,
Nascimento,L., Zutavern,T., McCombie,W.R. and Martienssen,R.A.
Genomic shotgun sequences from Zea mays (unfiltered)
Unpublished (2003)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8894
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ih28 row: a column: 11
Seq primer: -21M13UnivRev
Class: shotgun
High quality sequence stop: 787.

FEATURES
source Location/Qualifiers
1..787
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ih28all"
/lab_host="DH10b or JM107MA2"
/clone_lib="WGS-ZmaysU (DH10b or JM107MA2)"
/note="Organ: immature ears; Site: 1: Xba I; Site 2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector (.x/y
reads in M13mp19, .b/g reads in pUC19). The ligation was

ORIGIN

transformed into DH10b or JM107MA2."

Query Match 4.1%; Score 87.6; DB 29; Length 787;
Best Local Similarity 51.6%; Pred. No. 2e-11;
Matches 252; Conservative 0; Mismatches 229; Indels 7; Gaps 2;

QY 4 TCCCAACTTTTAGGAATGATCTTAAATTTTAGTTTAAAGTTCAAGTTAGAAAATCT 63
|||
Db 260 TCATATTTCTCGGAAGTGAATCTAAGGAACCTTGTGTGAATCCACATCCGGTACATTA 319
|||
QY 64 TTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAAACG-GTGTACATCTCTCCGATG 122
|||
Db 320 AACCAAGTCTTTGAGTTCATTCTCTCATTTCAACCGATTGAACATATCGATACA 379
|||
QY 123 GACTCACTCGTTCATTCGAAAGTTGGAAGAGTGATAGAAATATTCATTTTGGAT 182
|||
Db 380 CTCTCATCTCTTTTCATAATGAATGTCTCAAACTTCCTTTGCAACATATAGTTTGGCA 439
|||
QY 183 TCTTCACTCGTTCGTTGGTTCATGAGTGACCTCAAGAGTCTCCCAATATCAAAAGCC 242
|||
Db 440 CTCCTTCAAAATGTAGTTCCTTCATGAATTTCCATTTAGTTTATCCAAATCTCATGAGCG 499
|||
QY 243 GAATCACAAATTTGAATGTGATTTGAATTCATTTTGTCTAATGCACAAAAACAGGGCATTC 302
|||
Db 500 GTTGTGAGATTTTGTATACGATTGAATCTATATATCAAGAGCATCATAAAGACATTC 559
|||
QY 303 ATAGCCTTTGTTTAAAGCAAAACATTTCTTCCTCCGATTCATCCCATTCGCTCATCGGA 362
|||
Db 560 ATGCTTTGATCATTTTGTGAGGATATTTCTCATTAACGAGTGGTTTCATCTTC--- 616
|||
QY 363 AGAGAAAATTTTGAATTCATTTTCACAAATAGACCAAGCTCGAAATCCCATGGAATG 422
|||
Db 617 ---TTCATACACAAATCCATCCCTAACAAATTTGGCCAAATTTTTCACCCATTTGCTCTA 673
|||
QY 423 AGGAAGATCTCTCATATGATTTTCCAAATACATATGTAATTCGACTCATTTAAACATAGTGGGA 482
|||
Db 674 AGATGATTCGATCTCTTATTTTCCATAATCAATAATTTTCCATCAAGTGGGTGGC 733
|||
QY 483 TGTGTAAT 490
|||
Db 734 TTCCCAAT 741
|||

RESULT 9
BZ778766/c BZ778766 608 bp DNA linear GSS 14-MAR-2003
LOCUS ih84d10.b1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone
DEFINITION ih84d10, genomic survey sequence.

ACCESSION BZ778766
VERSION BZ778766.1 GI:28956209
KEYWORDS GSS.

SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 608)
Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
Katzenburger,F., King,L., Miller,B., Muller,S.,
Nascimento,L., Zutavern,T., McCombie,W.R. and Martienssen,R.A.
Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ih84 row: d column: 10
Seq primer: -21M13UnivFwd
Class: shotgun

REFERENCE AUTHORS
TITLE
JOURNAL
COMMENT

High quality sequence stop: 608.

FEATURES

source
1. .608
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ih84d10"
/lab_host="DH5a"
/clone_lib="WGS-ZmaysF (DH5a methyl filtered)"
/note="Organ: Immature ears; Site 1: Xba I; Site 2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector (x/y
reads in M13mp19, b/g reads in pUC19). The same ligation
was transformed into DH5a."

ORIGIN

Query Match 4.1%; Score 87.4; DB 28; Length 608;
Best Local Similarity 50.9%; Pred. No. 2.2e-11;
Matches 259; Conservative 1; Mismatches 242; Indels 7; Gaps 2;
QY 4 TCCCAACTTTTAGGAATGGATCTTAAATTTTAGTTTAAAGTTCAAGTTAGAAAAATCT 63
Db 512 TCATCTTATCCGGAAGAGATCTAAGAACTATGAGTGAATCCACATCTGGTACATTA 453
QY 64 TTACCAAGAGCTTTGAGTCATTTGATGACATCCGCGTGAACGG-TGTACATGTCCTCGATG 122
Db 452 AAACCAAGTCCCTTGAGTTCACTCACTCTCACTTTAACCGGTTGAACATATCAGATACA 393
QY 123 GACTCACTGGTTTCATTCGGAAGAGTTCGAAGAGTGCATAAGATATGATTTGGAT 182
Db 392 CTTTCATCTCTTTCATATGAATGCTCAAAATTTTCTTGCACATATAGCTTGGCA 333
QY 183 TCTTTCACATCGGTTGGTGCTTCATCAGTGACCTCAAGAGTCTCCAAATATCAAAAGCC 242
Db 332 CTTTTCACAAATGTAGTTCCTTCATGAAATTCATAGCTTAATCCAAATCICATGAGG 273
QY 243 GAATCAAAATGAATGTGATTAATTCATTTTCTTAATGCACAAACACAGGGGATTC 302
Db 272 GTTGTGAATCTTGTATGACGGTTGAATCACTTATATCAAGAGCATCATAGAATACATTC 213
QY 303 ATAGCCTTGTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATCGCTCATCGGA 362
Db 212 ATGCTGTGATCATTTGAAGATGTTTTCACTGTCACCAACAGATGGTTCACTCTTCTTC 153
QY 363 AGAGAAAAATTTTGAATCCATTTTCGACAAATAGACCAAGCTCGAAATCCATGGAAATG 422
Db 152 BACACAA-----CAAAACCATCTTGACAAATTTGGCCAAATTTCCACCCATGTCTTA 99
QY 423 AGGAAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACTCATATAACATAGTGGGA 482
Db 98 AGATGATTAACATCTTATTTTCCAAATAGTCGTAGTTGTTCCCATCAAAATCCGCTGGC 39
QY 483 TGTGTAATGAATGACCCCATGCSCTAT 511
Db 38 TTCCCAATGTGCATTTCTCGAGGCGTCT 10

RESULT 10

CG795738
LOCUS
DEFINITION
ZMMBBb0329L23.f ZMMBBb Zea mays subsp. mays genomic clone
ZMMBBb0329L23 5', genomic survey sequence.
ACCESSION
CG795738
VERSION
CG795738.1 GI:38212973
KEYWORDS
GSS.
SOURCE
Zea mays subsp. mays (maize)
ORGANISM
Zea mays subsp. mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

REFERENCE

1. (bases 1 to 767)
AUTHORS
Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.,
and Wing, R.
TITLE
Sequencing of the maize genome
JOURNAL
Unpublished (2003)
COMMENT
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0329 row: L column: 23
Seq primer: T7
Class: BAC ends.

FEATURES

Location/Qualifiers
1. .767
/organism="Zea mays subsp. mays"
/mol_type="genomic DNA"
/cultivar="B73"
/sub_species="mays"
/db_xref="taxon:4578"
/clone="ZMMBBb0329L23"
/lab_host="DH10B"
/clone_lib="ZMMBBb"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Zea mays L. ssp. mays"

ORIGIN

Query Match 4.0%; Score 87.2; DB 29; Length 767;
Best Local Similarity 52.9%; Pred. No. 2.5e-11;
Matches 258; Conservative 0; Mismatches 223; Indels 7; Gaps 3;
QY 7 CAACITTTAGGAATGGATCTTAAATTTTAGTTTAAAGTTCAAGTTAGAAAAATCTTTTA 66
Db 158 CATTTCTCTGGAAGGGATCTAAGGAACCTTATGAGTAAATCCACATCCGGTACATCAAGA 217
QY 67 CCAAGAGCTTTGAGTCCATTTGATGACATCCGCGAAACGG-TGTACATGTCCTCGATGGAC 125
Db 218 CCAATATCCCTTGAGCTCATTTACTATCTCATTTAACCGTTGACATGACAGTACATT 277
QY 126 TCATTTGGTTTCATTCGGAAAAAGTTGAAAGAGTGCATAAGAAATATTGATTTTGGATTCT 185
Db 278 TCAATTTCTTCATAATAAATTTGCTCAATTTTCCCTTCGCACACATATAATTTGACACTC 337
QY 186 TTCACCTCGGTTGGTGGCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAGCGGA 245
Db 338 TTCACAAATTTGATGTTCCCTTCATGAAATTTCCATTTAGCTTGCATCAAAATATCATAAAGCGGTT 397
QY 246 TCACAAATTTGAATGTGATTTGATTTTGTCTAATGCACAAACACAGGCGATTCATA 305
Db 398 GTGAGATCTTGATCTTATTTGAACTCAGTGTCTAGAGCATCATAGACATTCATG 457
QY 306 GCGTTTGTGTTTAAAGCAAAAACATTTCTTCTCCGATTCATCCCATTCGCTCATCGGAAGA 365
Db 458 GCTTGACCATTTAGTAACATTTTTCATTATC---ACTAGTCAATAGTTTCATC--TTGC 511
QY 366 GAAATTTTGAATTCATTTTCGACATAGACCAAGCTCGAAATCCATGGAATGAGG 425
Db 512 TTCAATACAAACAAATTCATCTCTGGCAATTTGGCAAAACCTTTCCACCCATTCCTCTATGA 571
QY 426 AAGATCCTCATATGAGTTTTCAAATACATGTAATTCGACTCATTAACATAGTGTGATGT 485
Db 572 TGCATTCACATCTTATTTTCCAAATAGTCATTAATTTGTTCCCATCAAAATGTGATGCTTT 631
QY 486 GTAATGAA 493
Db 632 CCAATGCA 639

```
RESULT 11
CG227901
LOCUS
DEFINITION
  CG227901 873 bp DNA linear GSS 22-AUG-2003
  genomic survey sequence.
ACCESSION
CG227901
VERSION
CG227901.1 GI:34127787
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 873)
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
  Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other_GSSs: OG0BU24TV
  Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
    source
    1..873
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBMA0682D23"
    /notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"
ORIGIN
Query Match 4.0%; Score 87; DB 29; Length 873;
Best Local Similarity 51.5%; Pred. No. 2.8e-11;
Matches 252; Conservative 0; Mismatches 230; Indels 7; Gaps 2;
QY 4 TCCCACTTTAGGAATGATCTTAAATTTTGTATTAAGTTCAGAGTTCGGAATCT 63
DB 21 TCATATTTCTCGAAGTGTCTAAGGAACCTTGTAGTGAATCCACATCCGGTACATTA 80
QY 64 TTACCAAGAGCTTTGAGTCCATTCATGACATCCGTGAACGG-TGTACATGTCCTCCGATG 122
DB 81 AAACCAAGTCCCTTGAGTTCATTCACATCTCATTTTAAACCGGTTGAACATATCGGATACA 140
QY 123 GACTCACTTGGTTTCATTCGGAAGAGTTTTCGAAAGAGTGCATAGAAATATTGATTTGGAT 182
DB 141 CTCTCATCTCTTTCATATAAATGCTCAAACCTTCCCTTTGCACACATATAGTTTGCA 200
QY 183 TCCTTCACTCGGTTGGTTCCTTATGAGTGCACCTCAAGAGTCCCTCCAAATATCAAAAGCC 242
DB 201 CTCTTCACAATTTAGTGTCTCTCATGAATTTCCATTTAGCTTAATFCCAAATCTCATGAGCA 260
QY 243 GAATCACAAAATGAAATGTGATTAATTTGCTAATGCACAAAACAGGGCAATTC 302
DB 261 GTTGTGAGATCTTGATGCGATTTGAATCACTTATATATCAAGAGCATCATGAGACATTC 320
QY 303 ATAGCCCTTTGGTTTAAAGCAAAACATTTCTTCGGAATTCATCCCATTCGCTCATCGGA 362
DB 321 ATGGCTTGATCAATTTGTAAGGATATCTCATTTGTCACTAACCGGATGTTTCATCTCTTC 380
QY 363 AGAGAAATTTTGAATTCATTTTCGACAAATAGACCAAGCTCGAATCCATGGAATG 422
DB 381 AACACAA-----CAATTCATCCCTGACAAATGGCCAAATTTTCCACCCCATTCGCTTTA 434
```

```
QY 423 AGGAAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACTCATTAACATAGGTGGA 482
DB 435 AGATGCATTGACATCTTATTTTCTAATAGTCATAGTTATTTCTCATCAAAATCGGTGGC 494
QY 483 TGTGTAATG 491
DB 495 TTCCCAATG 503

RESULT 12
BZ527252
LOCUS
DEFINITION
  BZ527252 707 bp DNA linear GSS 16-DEC-2002
  OGAI960TC ZM2_0.7_1.5_KB Zea mays genomic clone ZMMBMA0069J24,
  genomic survey sequence.
ACCESSION
BZ527252
VERSION
BZ527252.1 GI:27067787
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 707)
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
  Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
    source
    1..707
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBMA0069J24"
    /clone_lib="ZM2_0.7_1.5_KB"
    /notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"
ORIGIN
Query Match 4.0%; Score 86.6; DB 28; Length 707;
Best Local Similarity 52.8%; Pred. No. 3.5e-11;
Matches 259; Conservative 0; Mismatches 224; Indels 7; Gaps 3;
QY 4 TCCCACTTTAGGAATGATCTTAAATTTTGTATTAAGTTCAGAGTTCAGAAATCT 63
DB 120 TCATCTTCTCTGGAAGAGATCTAAGGAACCTTATGAGTGAATCCACATCAGGTACATTA 179
QY 64 TTACCAAGAGCTTTGAGTCCATTCATGATGACATCCGTGAACGG-TGTACATGTCCTCGATG 122
DB 180 AAACCAAGTCCCTTGAGTTCATTCATCTCATCTATCTATTAAACCGGTTGAACATATCAGATACA 239
QY 123 GACTCACTTGGTTTCATTCGGAAGAGTTTTCGAAAGAGTGCATAGAAATATTGATTTGGAT 182
DB 240 CTTTCACTTATTTCAATGAATTTGCTCAATTTTCCCTTTGCACACATATAGCTTTGGCA 299
QY 183 TCCTTCACTCGGTTGGTTCCTTCATGAGTGACCTCAAGAGTTCCTCAAAATATCAAAAGCC 242
DB 300 CTCTTCATAATTTAGTTCCTTCATGAATTTCCATTTAGCTTAATFCCAAATCTCATGAGCA 359
QY 243 GAATCACAAAATGAAATGTGATTAATTTTGTCTAATGCACAAAACAGGGCAATTC 302
DB 360 GTTGTGAGATCTTGTGATACGAATTTGAATCTCATTTATGTCTCAAGAGCATCATGAAATACATTC 419
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 14:23:59 ; Search time 8134.53 Seconds
(without alignments)
11509.077 Million cell updates/sec

Title: US-09-892-635A-45
Perfect score: 2160
Sequence: 1 ggatcccaacttttaggaat.....agctctcattcaccatgg 2160

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_btg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2110	97.7	4924	6	BD080859	Banana pr
2	185.2	8.6	298	8	AF399940	Musa acumin
3	122.2	5.7	282699	8	AF506028	AF506028 Poncirus
4	87.8	4.1	86019	2	AC133571	AC133571 Medicago
5	87.8	4.1	121952	8	AC146745	AC146745 Medicago
6	83.8	3.9	192284	2	AC146910	AC146910 Medicago
7	83.4	3.9	132605	8	AP004957	AP004957 Lotus cor
8	83.2	3.9	141590	2	AC126788	AC126788 Medicago
9	83	3.8	141590	2	AC126788	AC126788 Medicago
10	81.4	3.8	125883	8	AC124959	AC124959 Medicago
11	81.2	3.8	100810	8	AF411804	AF411804 Lycopersi
12	79.8	3.7	196632	8	AP004500	AP004500 Lotus cor
13	79.4	3.7	99606	8	AC116033	AC116033 Zea mays
14	78.8	3.6	129921	8	AC146568	AC146568 Medicago
15	78.4	3.6	135983	2	AC144538	AC144538 Medicago
16	78.4	3.6	136470	2	AC145222	AC145222 Medicago
17	76.6	3.5	97579	2	AC146722	AC146722 Medicago
18	74.4	3.4	93815	2	AC137663	AC137663 Cicer ari
19	73.6	3.4	133116	2	AC146631	AC146631 Medicago
20	73.4	3.4	809	6	A70187	A70187 Sequence 48
21	71.8	3.3	104684	8	AP006124	AP006124 Lotus cor
22	71.6	3.3	107979	8	AP006146	AP006146 Lotus cor
23	70.6	3.3	127857	2	AC133780	AC133780 Medicago
24	70.4	3.3	80930	8	AP004894	AP004894 Lotus cor
25	70.4	3.3	127739	2	AC145156	AC145156 Medicago
26	70.2	3.2	59202	8	AF220603	AF220603 Lycopersi
27	69.6	3.2	167484	2	AC145225	AC145225 Zea mays
28	69	3.2	101110	8	AP006141	AP006141 Lotus cor
29	69	3.2	114097	2	AC122161	AC122161 Medicago
30	68.8	3.2	75051	8	AP006114	AP006114 Lotus cor
31	68.8	3.2	112721	8	AP006401	AP006401 Lotus cor
32	68.8	3.2	127049	8	AP006428	AP006428 Lotus cor
33	68.8	3.2	129532	8	AP004482	AP004482 Lotus cor
34	68	3.1	72616	8	AP006091	AP006091 Lotus cor
35	68	3.1	104321	8	AP006142	AP006142 Lotus cor
36	67.4	3.1	105246	8	AC139747	AC139747 Medicago
37	67.4	3.1	106730	8	AC144724	AC144724 Medicago
38	67.4	3.1	115995	8	AC146566	AC146566 Medicago
39	67.2	3.1	79882	8	AP006118	AP006118 Lotus cor
40	67.2	3.1	80018	8	AP006355	AP006355 Lotus cor
41	67.2	3.1	86350	8	AP006139	AP006139 Lotus cor
42	67.2	3.1	98492	8	AP006085	AP006085 Lotus cor
43	67.2	3.1	100864	8	AP006391	AP006391 Lotus cor
44	67.2	3.1	106033	8	AP006128	AP006128 Lotus cor
45	67.2	3.1	108987	8	AP004511	AP004511 Lotus cor

ALIGNMENTS

RESULT 1
BD080859
LOCUS
DEFINITION Banana proteins, DNA, and DNA regulatory elements associated with
fruit development.
ACCESSION BD080859
VERSION BD080859.1 GI:22626462
KEYWORDS JP 2001517446-A/9.
SOURCE Musa acuminata
ORGANISM Musa acuminata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Zingiberales; Musaceae;
Musa.
REFERENCE 1 (bases 1 to 4924)
AUTHORS May, G. and Clendennen, S.

Pred. No. is the number of results predicted by chance to have a

TITLE Banana proteins, DNA, and DNA regulatory elements associated with fruit development
JOURNAL Patent: JP 2001517446-A 9 OCT-2001;
BOYCE THOMPSON INSTITUTE FOR PLANT RESEARCH INC
COMMENT OS Musa acuminata (banana)
PN JP 2001517446-A/9
PD 09-OCT-2001
PF 23-SEP-1998 JP 2000512960
PR 25-SEP-1997 US 60/060062
PI GREGORY MAY,STEPHANIE CUENENEN
PC C12N15/09,A01H5/00,C07K14/415,C07K14/42,C07K14/825,C12N5/10,
PC C12N9/08,
PC C12N9/10,C12N9/24,C12N15/00,C12N5/00
CC Banana proteins, DNA, and DNA regulatory elements associated with fruit development
CC development
CC Key
FH Key
FT source
FT Location/Qualifiers
FT Location/Qualifiers
FT 1..4924
FT /organism='Musa acuminata (banana)'.
FT /organism='Musa acuminata'
FT /mol_type='genomic DNA'
FT /db_xref='taxon:4641'
ORIGIN
source
Query Match 97.7%; Score 2110; DB 6; Length 4924;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 1; Indels 9; Gaps 3;
QY 1 GGATCCCAACTTTTAGGAATGATCTTAAATTTTAGTTATAAGTTCAAAGTTAGAAAA 60
DB 1 GGATCCCAACTTTTAGGAATGATCTTAAATTTTAGTTATAAGTTCAAAGTTAGAAAA 60
QY 61 TCTTTACCAAGAGCTTTGAGTCCATGTATGATACATCCGTGAAACCGGTGTACATGTC 120
DB 61 TCTTTACCAAGAGCTTTGAGTCCATGTATGATACATCCGTGAAACCGGTGTACATGTC 120
QY 121 TGGACTCACTGGTTTCAATCGGAAAGTTCCGAAGAGTGCATAGAATATGATTTGG 180
DB 121 TGGACTCACTGGTTTCAATCGGAAAGTTCCGAAGAGTGCATAGAATATGATTTGG 180
QY 181 ATCTTTCACTCGGTGGTCCCTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAAG 240
DB 181 ATCTTTCACTCGGTGGTCCCTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAAG 240
QY 241 CCGAATCAAAATGAATGTGAATTCATTTTGTCTAATGCAACAAACAGGGCAT 300
DB 241 CCGAATCAAAATGAATGTGAATTCATTTTGTCTAATGCAACAAACAGGGCAT 300
QY 301 TCATAGCCTTTGTGTTTAAAGCAAAACATTTCTCCGATTCATCCCATTCGCTCATCG 360
DB 301 TCATAGCCTTTGTGTTTAAAGCAAAACATTTCTCCGATTCATCCCATTCGCTCATCG 360
QY 361 GAAGAGAAATTTTGAATTCATTTTCGAATAGACCAAGCTCGAAATCCATGCAATG 420
DB 361 GAAGAGAAATTTTGAATTCATTTTCGAATAGACCAAGCTCGAAATCCATGCAATG 416
QY 421 GAAATGAGAGAGATCCTCATATGAGTTTCCAAATACATGTAATTCGATCTCAATTAACATA 480
DB 417 GAAATGAGAGAGATCCTCATATGAGTTTCCAAATACATGTAATTCGATCTCAATTAACATA 476
QY 481 GGTGATGTGAATGAATGACCTCATGCTCATCTCTCTGGTATTAAACCAATAT 540
DB 477 GGTGATGTGAATGAATGACCTCATGCTCATCTCTCTGGTATTAAACCAATAT 536
QY 541 GAGAGTGAGCCTTGCTCTGATACCAATTTGTAGGATCAGAGTGGCACTAAGAGAGGGGG 600
DB 537 GAGAGTGAGCCTTGCTCTGATACCAATTTGTAGGATCAGAGTGGCACTAAGAGAGGGGG 596
QY 601 GAGTGAATAGTGAGTGAATTAACATATTAAGTTTAAATTAAGTTCGTAATACGAG 660
DB 597 GAGTGAATAGTGAGTGAATTAACATATTAAGTTTAAATTAAGTTCGTAATACGAG 656

QY 661 AAGATTTCGTTTAAATAGTAACTTGAAGTAGTGAATGAAACCAAAAGTTAACAGTAGTGTAAA 720
DB 657 AAGATTTCGTTTAAATAGTAACTTGAAGTAGTGAATGAAACCAAAAGTTAACAGTAGTGTAAA 716
QY 721 TAACAATTTTCGGGAAAGTAAGAACTCACACATTCAGAGAAATACCAATTTTAAAGTGTGT 780
DB 717 TAACAATTTTCGGGAAAGTAAGAACTCACACATTCAGAGAAATACCAATTTTAAAGTGTGT 776
QY 781 CGGTCAAAATGACCTACATCATTGTAAGCCCTTTCTGGAAGAGGTCCCAATTCAC 840
DB 777 CGGTCAAAATGACCTACATCATTGTAAGCCCTTTCTGGAAGAGGTCCCAATTCAC 836
QY 841 TAGCAAAATCACTTTGAAGGGAAGGACAAATACCTCTCTTACNACCTTTTCAATGGTTC 900
DB 837 TAGCAAAATCACTTTGAAGGGAAGGACAAATACCTCTCTTACNACCTTTTCAATGGTTC 896
QY 901 ATACTCTTTACAAATTTTCAACGAGAAAGAGAGGTGAACATGCAAGCAATTTGAAACAA 960
DB 897 ATACTCTTTACAAATTTTCAACGAGAAAGAGAGGTGAACATGCAAGCAATTTGAAACAA 956
QY 961 GACTTGCTTAAAGACTTTGCTTAAGGCTTTTCTCAATCTATTGCTTCTCAAAAGTGTGA 1020
DB 957 GACTTGCTTAAAGACTTTGCTTAAGGCTTTTCTCAATCTATTGCTTCTCAAAAGTGTGA 1016
QY 1021 TTCTCTGCTGAGAAATGAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGCTCCA 1080
DB 1017 TTCTCTGCTGAGAAATGAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGCTCCA 1076
QY 1081 AATTTCAAAATCTCTGCGAGGTTCCCGAGGTTCCCGGTGCCACCGCTGTCACTGTTTGACA 1140
DB 1077 AATTTCAAAATCTCTGCGAGGTTCCCGAGGTTCCCGGTGCCACCGCTGTCACTGTTTGACA 1136
QY 1141 CTGACAGTGTACTAGAGGTGCCACCGCGAGACCTCTCGGTGTGGGCGGTGCCACCGC 1200
DB 1137 CTGACAGTGTACTAGAGGTGCCACCGCGAGACCTCTCGGTGTGGGCGGTGCCACCGC 1196
QY 1201 CTAGACTTTTTCAGCTCACTGTTGGATTCCAAACTTGACCCCAACCACTCGGAACCTCGG 1260
DB 1197 CTAGACTTTTTCAGCTCACTGTTGGATTCCAAACTTGACCCCAACCACTCGGAACCTCGG 1256
QY 1261 GTCCAAATGACCCGTAACCGGATTTATAGGATTAACCCCTTAATTCCTAACCCCTAATATG 1320
DB 1257 GTCCAAATGACCCGTAACCGGATTTATAGGATTAACCCCTTAATTCCTAACCCCTAATATG 1316
QY 1321 CAAACTAGCAACTGAAATATAGTCTTAAGCAAGTTTAAACCGGCAAACTCGAGTCT 1380
DB 1317 CAAACTAGCAACTGAAATATAGTCTTAAGCAAGTTTAAACCGGCAAACTCGAGTCT 1376
QY 1381 TCTTCCGCGCATCTTTGCGCAGACTCTGATATACCTTTGGATTCTTCTAGCGGACTCC 1440
DB 1377 TCTTCCGCGCATCTTTGCGCAGACTCTGATATACCTTTGGATTCTTCTAGCGGACTCC 1436
QY 1441 TAGTAGGTCCTCGATCTTTGCGGAGTTTAGCGAGTAGCGAGAACCTTCTCGGTGATCTCC 1500
DB 1437 TAGTAGGTCCTCGATCTTTGCGGAGTTTAGCGAGTAGCGAGAACCTTCTCGGTGATCTCC 1496
QY 1501 GCAACCGCGCATGATCTTTGCGGAGACTTTTGAAGACTTTGCAAGTCCCGATTTCT 1560
DB 1497 GCAACCGCGCATGATCTTTGCGGAGACTTTTGAAGACTTTGCAAGTCCCGATTTCT 1556
QY 1561 TCTCGTTTGGTTCCGACAGCATCTTAACGAAACTTCGGACTCTTGAATGATCCATCGAA 1620
DB 1557 TCTCGTTTGGTTCCGACAGCATCTTAACGAAACTTCGGACTCTTGAATGATCCATCGAA 1616
QY 1621 CTGACTCCGCTAGGCTTGTATATTTTCAAGGTATCAGTAACTTCTACATACCTA 1680
DB 1617 CTGACTCCGCTAGGCTTGTATATTTTCAAGGTATCAGTAACTTCTACATACCTA 1676
QY 1681 ACTCAATATAGTGAATTAATTAACCATCAATGATTTCATCATCAAAATTCGAC 1740
DB 1677 ACTCAATATAGTGAATTAATTAACCATCAATGATTTCATCATCAAAATTCGAC 1736

```

QY 1741 ATTCAACAAATCCGTAATCAATACCCATCAGGCTATAGTTACGTGACTTACTCTGT 1800
Db 1737 ATTCAACAAATCCGTAATCAATACCCATCAGGCTATAGTTACGTGACTTACTCTGT 1796
QY 1801 GATCCGTAAGTGAAGTTAGCGAGTCAATGATCCAGGTGCTGTCACCTATTGGCCGAAACACG 1860
Db 1797 GATCCGTAAGTGAAGTTAGCGAGTCAATGATCCAGGTGCTGTCACCTATTGGCCGAAACACG 1856
QY 1861 TATCCCTATCCAAATCCAGTCTTCTCAACTCTTCTAGCCCTACCGTCTCTTTTATT 1920
Db 1857 TATCCCTATCCAAATCCAGTCTTCTCAACTCTTCTAGCCCTACCGTCTCTTTTATT 1916
QY 1921 ACTTTTGAAGAAATCAAAATCAAAACAGATACAAATAACACGCTGAGACACTGTGACAT 1980
Db 1917 ACTTTTGAAGAAATCAAAATCAAAACAGATACAAATAACACGCTGAGACACTGTGACAT 1976
QY 1981 GCTAGTCTCTGGAAGCAATTAATTCGGCGATCCACAGACGTCGTGACGCTTCATCACCAC 2040
Db 1977 GCTAGTCTCTGGAAGCAATTAATTCGGCGATCCACAGACGTCGTGACGCTTCATCACCAC 2036
QY 2041 TTTTCTCTACAT-ACGATGTCATGCTTGTGTGATGACAGACACCAAGCTTGCCCT 2099
Db 2037 TTTTCTCTACAT-ACGATGTCATGCTTGTGTGATGACAGACACCAAGCTTGCCCT 2096
QY 2100 TTGTTGTGCTTAAACAGAGAGAGAGAGAGACAGACCGGATAGCTCTCTCATTCACCATG 2159
Db 2097 TTGTTGTGCTTAAACAGAGAGAGAGAGAGACAGACCGGATAGCTCTCTCATTCACCATG 2152
QY 2160 G 2160
Db 2153 G 2153

RESULT 2
AF399940/c 298 bp DNA linear PLN 09-OCT-2003
LOCUS Musa acuminata clone Radkal2 repeat sequence.
DEFINITION AF399940
ACCESSION AF399940
VERSION AF399940.1 GI:15529477
KEYWORDS
SOURCE Musa acuminata
ORGANISM Musa acuminata
REFERENCE 1 (bases 1 to 298)
AUTHORS Valarik,M., Sinkova,H., Hribova,E., Safar,J., Dolezelova,M. and Dolezel,J.
TITLE Isolation, characterization and chromosome localization of repetitive DNA sequences in bananas (Musa spp.)
JOURNAL Chromosome Res. 10 (2), 89-100 (2002)
MEDLINE 21989232
PUBMED 11993938
REFERENCE 2 (bases 1 to 298)
AUTHORS Valarik,M., Sinkova,H., Hribova,E., Safar,J., Dolezelova,M. and Dolezel,J.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Laboratory of Molecular Cytogenetics and Cytometry, Institute of Experimental Botany, Sokolovska 6, Olomouc 772 00, Czech Republic
FEATURES
source
Location/Qualifiers
1. .298
/organism="Musa acuminata"
/mol_type="genomic DNA"
/cultivar="Pisang Mas"
/db_xref="taxon:4641"
/clone="Radkal2"
repeat_region 1. .298
ORIGIN
Query Match 8.6%; Score 185.2; DB 8; Length 298;
Best Local Similarity 89.8%; Pred. No. 1.3e-35;
Matches 221; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

```

```

QY 6 CCAACTTTTAGGAATGATCTTAAATTTTAGTTATTAAGTTCAAAGTTAGAAAATCTTT 65
Db 246 CCAACTTTTAGGAAGGATCTTAAAGATTTTAGTGACTAGTTCAAAATTTAGAAAAT- 188
QY 66 ACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAAC-GGTGTACATGTCCTCGATGGA 124
Db 187 ACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAACCGGGGTGACATGTCCTCGATGGA 128
QY 125 CTCACCTTGGTTTCATTCGGAAAAAGTTTCGAAAGAGTGCATAAGAAATATTGATTTTGGATTC 184
Db 127 CTCACCTTGGTTTCATTCGGAAAAAGTTTCGAAAGATGCACAAGAAATATTGATTTTGGATTC 68
QY 185 TTTCACTCGTTGGTCTTCATGAGTGACCTCAAGAGTCTCCTCAATATCAAAAGCCGA 244
Db 67 TTTCACTCGTCTAGTGCCTTCATGAGTGACCTCTAGAGTTCTCCAAATATCAAAAGCCGA 8
QY 245 ATCACA 250
Db 7 GTACA 2

RESULT 3
AF506028/c 282699 bp DNA linear PLN 21-FEB-2003
LOCUS Poncirus trifoliata citrus tristeza virus resistance gene locus, complete sequence.
DEFINITION AF506028
ACCESSION AF506028
VERSION AF506028.1 GI:24461847
KEYWORDS
SOURCE Poncirus trifoliata
ORGANISM Poncirus trifoliata
REFERENCE 1 (sites)
AUTHORS Yang,Z.-N., Ye,X.-R., Choi,S., Molina,J., Moonan,F., Wing,R.A., Roose,M.L. and Mirkov,T.E.
TITLE Construction of a 1.2-Mb contig including the citrus tristeza virus resistance gene locus using a bacterial artificial chromosome library of Poncirus trifoliata (L.) Raf
JOURNAL Genome 44 (3), 382-393 (2001)
MEDLINE 21337599
PUBMED 1144697
REFERENCE 2 (bases 1 to 282699)
AUTHORS Yang,Z.-N., Ye,X.-R., Molina,J., Roose,M.L. and Mirkov,T.E.
TITLE Sequence Analysis of a 282-Kilobase Region Surrounding the Citrus Tristeza Virus Resistance Gene (Ctv) Locus in Poncirus trifoliata L. Raf
JOURNAL Plant Physiol. 131 (2), 482-492 (2003)
MEDLINE 22474771
PUBMED 12586873
REFERENCE 3 (bases 1 to 282699)
AUTHORS Yang,Z.-N., Ye,X.-R., Molina,J., Roose,M.L. and Mirkov,T.E.
TITLE Direct Submission
JOURNAL Submitted (24-APR-2002) Dept. of Plant Pathology and Microbiology, The Texas A&M University Agricultural Experiment Station, 2415 E. Hwy 83, Weslaco, TX 78596, USA
FEATURES
source
Location/Qualifiers
1. .282699
/organism="Poncirus trifoliata"
/mol_type="genomic DNA"
/cultivar="Pomeroy"
/db_xref="taxon:37690"
/clone="derived from BAC clones 27A14, 20J24, 83D17, and 84F5"
gene <306. .>5092
/mRNA join(<306. .377,493. .534,931. .972,2377. .2499,2870. .2926, 3477. .3614,4240. .4281,4312. .4379,5057. .>5092)
/product="Ctv.1"
join(<306. .377,493. .534,931. .972,2377. .2499,2870. .2926,
CDS

```


AGGFFGKNPMLNGRGTIALLRDRAFFLNDKLLVGMRLCNVPSRKNLTLYSOP
EAIIPPYDCELNNGRSGSSDGTETPEKGGNEETIOSGDIHLPTREV
TPRLFFFTLHVSRHLFVYLTLFWSDSGLKLDDEPLVKVEIALNAPNCODLL
SPSLVESRLVDVYAGMDNKLIPMSMRKRAQVSGESSIAPFSQKSNTPSKTSAPAL
LNIIVNKASGVRLTHGNNCITE"

Query Match 5.7%; Score 122.2; DB 8; Length 282699;
Best Local Similarity 59.8%; Pred. No. 1.1e-19;
Matches 222; Conservative 0; Mismatches 148; Indels 1; Gaps 1;
QY 6 CCAACTTTAGGAATGATCTTAAATTTTAAAGTTCAAGTTAGAAAAATCTTT 65
Db CCACTCTCTAGTAGAGATCTAATGATTTCTTGACITTTTCGGCTATTGAAAAAGTCTT 72679
QY 66 ACCAAGAGCTTGGATCCATTGATCGATCCGTGAAC-GGTGTACATGCTCCGATGGA 124
Db TCCTAGGGCTCCTAGAGTGTTCTATGTCGGTGAATCTAGTATACATAGATACACATT 72619
QY 125 CTCACITGGTTTATTCGGAAGGTTGAAAGATGATGAATAATTTGTTGGATTC 184
Db TTCATTTTGTTCATTGMAACATTCGTATTGTCGAGTGATCTACTATTATTAGATTC 72559
QY 185 TTTCACCTGGTTGGTGGCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAGCCGA 244
Db TTTTCACTTGGTTCGTACCTTCATAGACAACCTCAAGTTTGTGCCAAATCTCATTAGCACT 72499
QY 245 ATCAAAATGAATGATGATTGAATTCATTTTCTCTAATGCAAAACAGGCGATTCAT 304
Db TTGCAACTAGACCTCTATGAACCTTTACTTATCTAGTGCACAGAACAGGCAATTCAT 72439
QY 305 AGCCTTTGTGTTTAAAGCAAAACATTTCTCCGATTCATCCCATTCGGTCTATCGGAAG 364
Db GGCTTGGAAATTAGAGAAGCTTTTCTCTATCGAATTCATTCAAATCCCGTGAGGTTT 72379
QY 365 AGAAATTTT 375
Db TGGAAATATCTT 72368

RESULT 4
AC133571/c 86019 bp DNA linear HTG 05-NOV-2003
LOCUS Medicago truncatula clone mth2-10h12, WORKING DRAFT SEQUENCE, 2
DEFINITION ordered pieces.

AC133571 28 GI:38175857
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.
KEYWORDS Medicago truncatula (barrel medic)
SOURCE Medicago truncatula
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE 1 (bases 1 to 86019)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE Medicago truncatula BAC Clone mth2-10h12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 86019)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (14-SEP-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 3 (bases 1 to 86019)
AUTHORS Shaull,S., Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B.,
Cook,D., Kim,D. and Roe,B.A.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT On Nov 5, 2003 this sequence version replaced gi:38044179.

----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 23532: contig of 23532 bp in length
* 23533 23632: gap of unknown length
* 23633 86019: contig of 62387 bp in length.

FEATURES
source

1..86019
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-10h12"
/clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match 4.1%; Score 87.8; DB 2; Length 86019;
Best Local Similarity 55.0%; Pred. No. 5.1e-11;
Matches 193; Conservative 0; Mismatches 157; Indels 1; Gaps 1;
QY 1 GGATCCCAACTTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAAAAGTTAGAAAAA 60
Db GGTCTCCATTGAGTTGGAAGACATCTTAGTATTCTTTCTTCTGTCGGTGAATAA 56956
QY 61 TCTTTACCAAGAGCTTTGAGTCCATTGATCAGATCCGTGAAACG-GGTACATGCTCCG 119
Db GTTTTCCAGTGATCTTAGTTCACTTACTATAGATGTAATCTAGATACATTTCAATCA 56896
QY 120 ATGGAATCACTTGGTTTCATTCGGAAGAGTTCCAAAGAGTGCATAAGAAATGATTGTTG 179
Db ATGTTTTCATCTTCAATCATTTCAAAGATTTCAAACCTTTCTCACCAATGTCAATCTTT 56836
QY 180 GATTCTTTCACCTCGTTGGTGGCTTCATGAGTCAAGTCAAGTCTCCGAAATATCAAA 239
Db GTTTCTTTAACATGGCTTGTCTCTCATGATGAGACTTTGAGAGATATCCCAACACTTTT 56776
QY 240 GCGAATCACAAAATTCGAAATGTGATTGAATTTTGTCTAATGCACAAAAACAGGCA 299
Db GCAGTATCACATTCATCACTCTTTTTCATCTTCTTCCATGCTTAGTGACATGATAGAA 56716
QY 300 TTCAATAGCCTTTGTTGTTTAAAGCAAAACATTTCTTCCGATTCATCCCAT 350
Db AGTCGAGCTTTAGAAATTTAAAGATATTCTTTCTTATCTCTTTTCTCCAT 56656

RESULT 5
AC146745

LOCUS Medicago truncatula clone mth2-53p19, complete sequence.
DEFINITION AC146745
ACCESSION AC146745
VERSION AC146745.17 GI:39573788
KEYWORDS HTG.

SOURCE
ORGANISM Medicago truncatula (barrel medic)
Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE 1 (bases 1 to 121952)
AUTHORS Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
and Roe,B.A.

TITLE Medicago truncatula BAC Clone mth2-53p19
JOURNAL Unpublished

```

REFERENCE
AUTHORS      2 (bases 1 to 121952)
              Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
              and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (24-SEP-2003) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
REFERENCE
AUTHORS      3 (bases 1 to 121952)
              Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
              and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (06-DEC-2003) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
REFERENCE
AUTHORS      4 (bases 1 to 121952)
              Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
              and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (09-DEC-2003) Department Of Chemistry And Biochemistry,
              The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
              OK 73019, USA
COMMENT      On Dec 9, 2003 this sequence version replaced gi:39229182.
              ----- Genome Center
              Center: Department Of Chemistry And Biochemistry
              The University Of Oklahoma
              Center code:UOKNOR
              -----
FEATURES
source       Location/Qualifiers
              1..121952
               /organism="Medicago truncatula"
               /mol_type="genomic DNA"
               /db_xref="taxon:3880"
               /clone="mth2-59121"
               /clone_lib="Medicago truncatula BAC library H2"
ORIGIN
Query Match      4.1%; Score 87.8; DB 8; Length 121952;
Best Local Similarity 55.0%; Pred. No. 5.2e-11;
Matches 193; Conservative 0; Mismatches 157; Indels 1; Gaps 1;
QY 1 GATCCCCAACITTTTAGGAATGATCTTAAATTTTAGTTATAGTTCAAAGTTAGAAAA 60
Db 94945 GGTCTCCATGCTCTGGAGACATCTTAAATTTTCTTATTCGATCATGATGGTATAA 95004
QY 61 TCTTTACCAAGAGCTTTGAGTCATGATGATCCGCGAAC-GGTGATCATGCTCGG 119
Db 95005 GTTTTCCAGTGATCTTAATTCATTCACAATAGATGTAATCTGGAGTACATTTTCATCA 95064
QY 120 ATGGACTCACTGGTTTCATTTCGAAAGATTCGAAAGAGTGCATTAAGATATGATTTTG 179
Db 95065 ATGTTCTCATCTTCGTTTCATTTCGAGATTTCAAATTTTCTTACCCATATCTATTCTT 95124
QY 180 GATTCCTTTCACTCGGTTGGTGCCTTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAA 239
Db 95125 GTTTCCTTTACGTGCTAGTTCCCTCATGATGAATTTTAGAGTGTCCCAAACTTCTTTT 95184
QY 240 GCCGAATCACAATTTGAATGATGATTCATTTTGTCTAAATGCACAAACAGGGCA 299
Db 95185 GCAGTAACACATTCATCACTCTTTTCACCTTTCTCCATGCTTAGTGGCGATGATAGAAG 95244
QY 300 TTCATAGCCTTTGTGTTTAAAGCAAAAACATTCCTTCGATTCATCCCAT 350
Db 95245 AGTCGAGCTTUGAGTTTGAAGTACTTTTCTTCTTATCTTCTTTTGACCAT 95295
RESULT 6
AC146910/c 109284 bp DNA linear HTG 06-DEC-2003
LOCUS Medicago truncatula clone mth2-59121, WORKING DRAFT SEQUENCE, 2
DEFINITION unordered pieces.
AC146910
AC146910.5 GI:39229867
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

```

SOURCE ORGANISM

Medicago truncatula (barrel medic)
 Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE AUTHORS

1 (bases 1 to 109284)
 Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
 and Roe,B.A.
 Medicago truncatula BAC Clone mth2-59121
 Unpublished

TITLE JOURNAL

2 (bases 1 to 109284)
 Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
 and Roe,B.A.
 Direct Submission

REFERENCE AUTHORS

Submitted (21-OCT-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

REFERENCE AUTHORS

3 (bases 1 to 109284)
 Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D.
 and Roe,B.A.
 Direct Submission

TITLE JOURNAL

Submitted (06-DEC-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA

COMMENT

On Dec 6, 2003 this sequence version replaced gi:38371848.
 ----- Genome Center
 Center: Department Of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:UOKNOR

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 30237: contig of 30237 bp in length
 * 30238 30337: gap of unknown length
 * 30338 109284: contig of 78947 bp in length.

FEATURES source

1..109284
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /db_xref="taxon:3880"
 /clone="mth2-59121"
 /clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match 3.9%; Score 83.8; DB 2; Length 109284;
 Best Local Similarity 54.5%; Pred. No. 5.3e-10;
 Matches 189; Conservative 0; Mismatches 157; Indels 1; Gaps 1;
 QY 7 CAACCTTTAGGAATGATCTTAAATTTTAGTTATAGTTCAAAGTTAGAAAAATCTTTA 66
 Db 1100 CAAGATCTGGAGACATCTTAAATTTTCTTCTTCTTTCATGATGTAATATCTTT 1041
 QY 67 CCAAGAGCTTTGAGTCCATTCATGATGATCCGCTGAAACGGTG-TACATGCTCCGATGGAC 125
 Db 1040 TCAAGAGAGCGTAACTCATTCATATGATAGTAAACCTTCCATACATTCATCTATGTT 981
 QY 126 TCACCTTGTTTCATTCGGAAGATTCGAAAGAGTGCATAAGAATATTCATTTTGGATCT 185
 Db 980 TCGGTTCTCTTCATTCGATAGTTCAAACCTTTCGAACACATGTCATCTTCTTCT 921
 QY 186 TTCACCTCGTTGGTGCCTTTCATGATGATGACCTCAAGAGTCTCCAAATATCAAAAGCCGAA 245
 Db 920 TTTACGTGATTTGGTCCCTTCATGATGATGATACCTTCAATGTGTCCCAATTTCTTTGGCATTT 861
 QY 246 TCACAAATGAAATGTGATTCATTTTGTCTAATGCAACAAACAGGCGATTCATA 305


```

Db      860 TTGCATTGTGTAAGTCTATCATATCTTCCCTGCTCAATGCCAGTGTAGAAATAGTTTA 801
QY      306 GCCTTTGTGTTTAAAGCAAAACATCTTCTCCGATTCATCCCATTC 352
Db      800 GCCTTTAAATTAATAACATGCGCTTACCTCTCTCTTCTGACCAGTC 754

RESULT 7
LOCUS   AP004957                132605 bp    DNA        linear    PLN 22-JUL-2003
DEFINITION
Lotus corniculatus var. japonicus genomic DNA, chromosome 2,
complete sequence.
ACCESSION
AP004957
VERSION  AP004957.1    GI:21907975
KEYWORDS
HTG.
SOURCE   Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM
Lotus corniculatus var. japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
Lotus.
REFERENCE
1 Kaneko, T., Nakamura, Y., Asamizu, E., Kato, T., Sato, S. and Tabata, S.
Structural Analysis of a Lotus japonicus Genome. I. Sequence
Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb
Regions of the Genome
Unpublished
JOURNAL
REFERENCE
2 (bases 1 to 132605)
AUTHORS
Nakamura, Y.
TITLE
Direct Submission
JOURNAL
Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail: yna@kazusa.or.jp,
URL: http://www.kazusa.or.jp, Tel: 81-438-52-3935,
Fax: 81-438-52-3934)
FEATURES
source
1. .132605
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="2"
/clone="LjT34H20"
/clone_lib="LjT library"
/note="TAC clone: TW0134-synonym: Lotus japonicus"
ORIGIN
Query Match 3.9%; Score 83.4; DB 8; Length 132605;
Best Local Similarity 53.7%; Pred. No. 6.7e-10;
Matches 194; Conservative 0; Mismatches 166; Indels 1; Gaps 1;
QY 1 GGATCCCACTTTTAGGAATGGATCTTAAATTTTAGTTATAGTTCAAGTTAGAAAA 60
Db 112738 GGATCCAGCTTCAGGAAGACCTCTAAGAATCCTCATACATGATCACCAGTTGTGTA 112797
QY 61 TCTTTACCAAGAGCTTTGAGTCCATGATGA-CATCGTGAAACGGTGATCATGTCTCG 119
Db 112798 CTTTATTAGGGGTCTAATTCAGCAAGAGTAACTGAATCTGAGAACATGCTTCA 112857
QY 120 ATGGAATCACTGGTTTCATTCGGAAGAGTTCGAAAGAGTGCATAGAATATTGTTTG 179
Db 112858 ATGGAATCACTGGATTCCTTTCGGAAGGATTCATCTTCCTGATCAAGGACGCTTT 112917
QY 180 GATTCCTTCACTGGTTGGTCCCTCATGATGACCTCAAGATCTCCAAATATCAAA 239
Db 112918 GACTCTTTCACCTCTTGTGTTTCATCATGGGTCTCTTCAAGAAATCAAGATACCTTTA 112977
QY 240 GCCGAATCAAAATGAAATGTCATTGAATTCATTTTGTCTAATGCACAAACAGGCA 299
Db 112978 GCATATCTCAGATCTGTAATCTTCTCGTATTCTTCAATTGAATACATCTGAGAAGATA 113037
QY 300 TTCATAGCCTTTGTGTTTAAAGCAAAACATCTCTCCGATTCATCCCATTCGCTCATC 359

```

```

Db      113038 GTTCTAGCTTTATGATGCAGTGAAGTAAAGCTCTCTTGCTTCGATCACTCTGTTTG 113097
QY      360 G 360
Db      113098 G 113098

RESULT 8
LOCUS   AC126788/c            141590 bp    DNA        linear    HTG 11-NOV-2003
DEFINITION
Medicago truncatula clone mth2-33121, WORKING DRAFT SEQUENCE, 8
unordered pieces.
ACCESSION
AC126788
VERSION  AC126788.9    GI:38257184
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Medicago truncatula (barrel medic)
ORGANISM
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 141590)
AUTHORS
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
TITLE
Medicago truncatula BAC Clone mth2-33121
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 141590)
AUTHORS
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
TITLE
Direct Submission
JOURNAL
Submitted (09-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
3 (bases 1 to 141590)
AUTHORS
Shauli, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
TITLE
Direct Submission
JOURNAL
Submitted (11-NOV-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Nov 11, 2003 this sequence version replaced gi:37059850.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code: UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 4380: contig of 4380 bp in length
* 4381 4380: gap of unknown length
* 4481 14258: contig of 9778 bp in length
* 14259 14358: gap of unknown length
* 14359 20669: contig of 6311 bp in length
* 20670 20769: gap of unknown length
* 20770 25083: contig of 4314 bp in length
* 25084 25183: gap of unknown length
* 25184 41406: contig of 16223 bp in length
* 41407 41506: gap of unknown length
* 41507 60351: contig of 18845 bp in length
* 60352 60451: gap of unknown length
* 60452 87167: contig of 26716 bp in length
* 87168 87267: gap of unknown length
* 87268 141590: contig of 54323 bp in length.
Location/Qualifiers
1. .141590
/organism="Medicago truncatula"
/mol_type="genomic DNA"
FEATURES
source
1. .141590
/organism="Medicago truncatula"
/mol_type="genomic DNA"

```


Medicago.
 1 (bases 1 to 125883)
 Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
 Cook, D., Kim, D. and Roe, B.A.
 Medicago truncatula BAC Clone mth2-23111
 Unpublished
 2 (bases 1 to 125883)
 Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
 Cook, D., Kim, D. and Roe, B.A.
 Direct Submission
 Submitted (06-JUN-2002) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 3 (bases 1 to 125883)
 Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
 Cook, D., Kim, D. and Roe, B.A.
 Direct Submission
 Submitted (05-MAR-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 4 (bases 1 to 125883)
 Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
 Cook, D., Kim, D. and Roe, B.A.
 Direct Submission
 Submitted (05-MAR-2003) Department Of Chemistry And Biochemistry,
 The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 On Mar 5, 2003 this sequence version replaced gi:24635962.

 Center: Genome Center
 Department Of Chemistry
 The University Of Oklahoma
 Center code:UOKNOR

FEATURES

source
 Location/Qualifiers
 1..125883
 /organism="Medicago truncatula"
 /mol_type="genomic DNA"
 /db_xref="taxon:3880"
 /clone="mth2-23111"
 /clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match 3.8%; Score 81.4; DB 8; Length 125883;
 Best Local Similarity 53.5%; Pred. No. 2.1e-09;
 Matches 192; Conservative 0; Mismatches 166; Indels 1; Gaps 1;
 1 GGATCCCACTTTTAGGAATGGATCTTAAATTTTGTATATAGTTCAAGTTAGAAAAA 60
 5238 GGTCTCCATACACTTGGGAAGACATCTTAATATTTTCTGATTTCTATCAATAGTAGAATA 5179
 61 TCTTTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAACG-GTGTACATGCTCCG 119
 5178 GCCTTTCCAGAGATCTCATTTCTATTCTATTGTGTAGTAATCTAGATACATCTCATCA 5119
 120 ATGGACTCATCTGGTTTCATTCGGAAAAAGTTCGAAAGAGTGCATAAGAAATATTGATTTTG 179
 5118 ATGGTTTCATTTTCACTCATTTTCAAAAACCTTCAAAATTTTCTGACTCCAAATATCAATCTCA 5059
 180 GATCTTTTCACTGGTGGTCCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAA 239
 5058 GTTCTTTTGACATGAGATGACCTCATGATATATTTTTCAGAGTGTCCCATCTCTTG 4999
 240 GCCGAATCAAAATGAAATGTGATTGATTTATTTTGTCTAATSCACAAACAGGGCA 299
 4998 GCATTTTACACTCATCACTCTTTCACCTTTCTCCATGGTTAGTGACATGATAGAAA 4939
 300 TTCATAGCCTTTGTGTTTAAAGCAAAAACATTTCTCCGATTCATCCCATTCGCTCAT 358
 4938 AGCTAGCTTTAGAGTTTAGTATACATGAGCCTTTTCACTGTTGACCATGCTTCTT 4880

RESULT 11
 AF411804

LOCUS
 DEFINITION
 AP411804 100810 bp DNA linear PLN 25-JUL-2002
 Lycopersicon esculentum BAC clone Clemson_ID 47113, complete
 sequence.
 ACCESSION
 VERSION
 AP411804.1 GI:15987769
 SOURCE
 ORGANISM
 Lycopersicon esculentum (tomato)
 Lycopersicon esculentum
 Sukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 100810)
 Van Der Hoeven, R., Ronning, C., Giovannoni, J., Martin, G. and
 Tanksley, S.
 Deductions about the Number, Organization, and Evolution of Genes
 in the Tomato Genome Based on Analysis of a Large Expressed
 Sequence Tag Collection and Selective Genomic Sequencing
 Plant Cell 14 (7), 1441-1456 (2002)
 12119366
 2 (bases 1 to 100810)
 van der Hoeven, R.S. and Tanksley, S.D.
 Direct Submission
 Submitted (20-AUG-2001) Plant Breeding, Cornell University, 252
 Emerson Hall, Ithaca, NY 14850, USA
 FEATURES
 source
 1..100810
 /organism="Lycopersicon esculentum"
 /mol_type="genomic DNA"
 /db_xref="taxon:4081"
 /clone="Clemson_ID 47113 BAC"
 ORIGIN
 Query Match 3.8%; Score 81.2; DB 8; Length 100810;
 Best Local Similarity 52.0%; Pred. No. 2.4e-09;
 Matches 205; Conservative 0; Mismatches 188; Indels 1; Gaps 1;
 4 TCCCACTTTTAGGAATGGATCTTAAATTTTGTATATAGTTCAAGTTAGAAAAATCT 63
 89844 TCCCAAGACCTTAGAAAAATCTTAAGTACTTTGTCAACCTGTTCTCAGTGGTATATACT 89903
 64 TTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAAC-GGTGTACATGCTCCGATG 122
 89904 TTACCAATGAGATTAGTTAGTTTATCAGATGCTGAATCTAGTATCATATCTGAGA 89963
 123 GACTCATCTGGTTTCATTCGAAAAAGTTCGAAAGAGTGCATAAGAAATATTGATTTGGAT 182
 89964 GATTTTCCGGAATTCATCTTAAAGGCTTCATATTCTGAACACAACTAGCAATCTTGAAT 90023
 183 TCTTTCACCTCGGTTGGTGCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAGCC 242
 90024 TTTCGACTTGAGTTGTGATCTTCATGAGAGTTTGTAAAGTATCCCATATTTGTTAGCA 90083
 243 GAATCAAAATGAAATGTGATTGAATTCATTTTGTCTAATGCACAAAACAGGCAATC 302
 90084 GAGGTAACTTCGATACGGTTAAACTCATCGGTCCTAGTCCACATACAGAATGCAC 90143
 303 ATAGCCTTTGTGTTAAACCAAAAACATTTCTTCGAGATCATCCCATTCGCTCATCGGA 362
 90144 TTAGCTTTTGCATTTCTTCTAGCATTTTATAATCGCTTCATTTATATGATCTGCAGGT 90203
 363 AGAGAAAAATTTTGAATCCATTTTCGACAATAG 396
 90204 TTAGGTACTTTTATTCCTTCACTATCAGTAATAG 90237
 RESULT 12
 AP004500/c
 LOCUS
 DEFINITION
 AP004500 106632 bp DNA linear PLN 22-JUL-2003
 Lotus corniculatus var. japonicus genomic DNA, chromosome 1,
 clone:LjT05P21, TM0032, complete sequence.
 ACCESSION
 VERSION
 AP004500.1 GI:17736867
 KEYWORDS
 HTG.

SOURCE Lotus corniculatus var. japonicus (Lotus japonicus)
ORGANISM Lotus corniculatus var. japonicus
TITLE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
Lotus.
REFERENCE 1
AUTHORS Sato,S., Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T. and Tabata,S.
TITLE Structural Analysis of a Lotus japonicus Genome. I. Sequence
Features and Mapping of Fifty-six TAC clones which cover the 5.4 Mb
Regions of the Genome
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 106632)
REFERENCE Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (13-DEC-2001) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
URL:htp://www.kazusa.or.jp, Tel:81-438-52-3935,
Fax:81-438-52-3934)
FEATURES
source
location/Qualifiers
1. .106632
/organism="Lotus corniculatus var. japonicus"
/mol_type="genomic DNA"
/variety="japonicus"
/db_xref="taxon:34305"
/chromosome="1"
/clone="LJ705P21"
/clone_lib="LJ7 library"
/note="TAC clone:TM0032-synonym: Lotus japonicus"
ORIGIN
Query Match 3.7%; Score 79.8; DB 8; Length 106632;
Best Local Similarity 53.9%; Pred. No. 5.4e-09;
Matches 185; Conservative 0; Mismatches 157; Indels 1; Gaps 1;
QY 11 TTTTAGGAATGATCTTAAATTTTAGTTATAAGTTCAGGTTAGAAAATCTTTACCAA 70
Db TCTGGGAGAGATCTCAAGATCTTCTTCAAGCTTCTCACTGACATAGTTCTCCCA 24508
QY 71 GAGCTTTGAGTCCATGATGATCATCGGTGAACG-GTGATACATGTCTCCGATGGACTCAC 129
Db AAGCAAAATGAACATGACAGATGACAGCAGCATGTGGAATCTGAGATAGTCTT 24448
QY 130 TTGGTTTCATTCGGGAAAAGTTCGAAAGATGCAATGAATATTTGATTTGGATCTTTCA 189
Db CTTTAGTTCATCATCAAAATTTTCAAAATTTGAGTGGTGAAGATCTGAAGTCTTGACATCTCTTA 24388
QY 190 CTCGGTTGGTGCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAGCCGAATCAC 249
Db CTCGAGTAGTTCCTTCATGTGAGTCTTCAGAAATTTCCAGATATCTTTCGCCACAGTAC 24328
QY 250 AAATTTGAAATGATGATGAATTCATTTTGTCTTAATGCAACAAACAGGGCATTCATAGCCT 309
Db ACGTATTGATCAGCTAAACATATCTTGTCAACTCCATTAAAGATAGCAATTCAGAGCT 24268
QY 310 TTGTGTTTAAGCAAAAACATCTTCTCCGATTCATCCCATTC 352
Db TTGAGTTTCCAAGACAGCTTCATCTTCAACACTGGTCCATTC 24225
RESULT 13
AC116033/c 99606 bp DNA linear PLN 04-OCT-2003
LOCUS
DEFINITION Zea mays genomic clone ZM15C05 sequence, complete sequence.
ACCESSION AC116033
VERSION AC116033.3 GI:37514986
KEYWORDS HTG
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 99606)
AUTHORS Nagaki,K., Song,J., Stupar,R.M., Parokorny,A.S., Yuan,Q.,
Qiyang,S., Liu,J., Hsiao,J., Jones,K.M., Dawe,K., Buell,R. and
Jiang,J.
TITLE Molecular and cytological analyses of large tracks of centromeric
DNA reveal the structure and evolutionary dynamics of maize
centromeres
JOURNAL Genetics 163 (2), 759-770 (2003)
MEDLINE 22503408
PUBMED 12618412
REFERENCE 2 (bases 1 to 99606)
AUTHORS Jiang,J., Nagaki,K., Yuan,Q., Vanaken,S., Utterbach,T.,
Gansberger,K. and Buell,R.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-2002) University of Wisconsin, Department of
Horticulture, Madison, WI 53706, USA
REFERENCE 3 (bases 1 to 99606)
AUTHORS Jiang,J.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2002) University of Wisconsin, Department of
Horticulture, Madison, WI 53706, USA
REFERENCE 4 (bases 1 to 99606)
AUTHORS Jiang,J.
TITLE Direct Submission
JOURNAL Submitted (28-NOV-2002) University of Wisconsin, Department of
Horticulture, Madison, WI 53706, USA
REFERENCE 5 (bases 1 to 99606)
AUTHORS Jiang,J.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-2002) University of Wisconsin, Department of
Horticulture, Madison, WI 53706, USA
REFERENCE 6 (bases 1 to 99606)
AUTHORS Jiang,J.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2003) University of Wisconsin, Department of
Horticulture, Madison, WI 53706, USA
COMMENT On Oct 4, 2003 this sequence version replaced gi:25141017.
FEATURES
source
location/Qualifiers
1. .99606
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
/chromosome="unknown"
/clone="ZM15C05"
/note="The BAC contains DNA sequences associated with
maize centromeres."
ORIGIN
Query Match 3.7%; Score 79.4; DB 8; Length 99606;
Best Local Similarity 51.9%; Pred. No. 6.8e-09;
Matches 256; Conservative 0; Mismatches 226; Indels 11; Gaps 3;
QY 4 TCCCAACTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAAGTTAGAAAATCT 63
Db TCATATTTCTTCGGAAGTGAATCTAAGAACTGTGTGTAATCCACATCCGGTACATTA 33064
QY 64 TTACCAAGAGCTTTGAGTCCATGATGACATCCGTTGAACCGT-CTACATGCTCTCCGATG 122
Db AACCAAGTCTTTGAGTTCATTTATCTCATTCATCAACCGATGAACATATCGGATACA 33004
QY 123 GACTCAGTCTGGTTTCATTTCGGAAAAGTTTCGAAAGAGTGCATAGAATAATTCATTTTGGAT 182
Db CTCATCTCTCTTCATATAATAATTTGCTCAAACTTCCCTTTGCACACATATAGTTTGCA 32944
QY 183 TCTTTCACTCGTTGGTGGCTTCATGATGACCTCAAGAGTCTCCCAATATCAAAAGCC 242
Db CTCCTCAATTTAGTTCCTTCATGAATTTCCATTTATCCAAATCTCATGAGCG 32884
QY 243 GAATCACAAATTTGAAATGTGATTCGAATTCATTTTGTCTAATGCACAAAACAGGGCATTC 302
Db GTTGTGAGATTTCTTGATACGATTTGACCATTATATCAAGAGCATCAATAAGACATTC 32824
QY 303 ATAGCCTTTGTGTTTAAAGCAAAACATTTCTTCGGAATTCATCCCAATCGCTCATCGGA 362

```

Db 32823 TTGGCTTGATCATTTGTGAGGATATTTCTCATTAATCACTAACAGATGGTTTCATCTTCCTTC 32764
QY 363 AGAGAAATTTTGAATCCATTTTTCGACATAGACMAAGCTGAATTCATGATGGA 422
Db 32763 AACACAA-----CAAAATCCATCCCTGACAAATGGCCAAATTT-----TTCCACCCATTCG 32714
QY 423 AATGAGGAAGATCCTCATATGAGTTTTCATATACATGTAATTCGACTCATTAACATAGG 482
Db 32713 TCTAAGATGATGACATCTTAATTTTCCAAATATATCATTAATTTTCATCAAAAGTGGCG 32654
QY 483 TGAATGTGTAATG 495
Db 32653 TGGCTTTCCAATG 32641

RESULT 14
AC146568 129921 bp DNA linear PLN 18-DEC-2003
LOCUS Medicago truncatula clone mth2-101018, complete sequence.
DEFINITION
AC146568
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 129921)
Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
Medicago truncatula BAC Clone mth2-101018
Unpublished
REFERENCE
2 (bases 1 to 129921)
Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
Direct Submission
Submitted (30-AUG-2003) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
3 (bases 1 to 129921)
Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
Direct Submission
Submitted (05-DEC-2003) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
4 (bases 1 to 129921)
Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D.
and Roe, B.A.
Direct Submission
Submitted (18-DEC-2003) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Dec 5, 2003 this sequence version replaced gi:38564350.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----

FEATURES
source
1. 129921
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-101018"
/clone_lib="Medicago truncatula BAC library H2"

ORIGIN
Query Match 3.6%; Score 78.8; DB 8; Length 129921;
Best Local Similarity 53.4%; Pred. No. 9.7e-09;
Matches 187; Conservative 0; Mismatches 162; Indels 1; Gaps 1;

```

```

QY 1 GGATCCCAACTTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAAAGTTAGAAAAA 60
Db 50577 GGCTGCCAAATCCCTTGTCATACATCTCAAAATTTTATAAGTGAATCTTGAATG 50636
QY 61 TCTTTACCAAGAGCTTTGAGTCCATTTGATGACATCCGTTGAAC-GGTGTATATGTCCTCG 119
Db 50637 TCATTACTAGGATTTGCAATTTTAAACAATATGGGTGAATCTCATTTGCACGAGGAG 50696
QY 120 ATGGACTCACCTTGGTTTTCATTCGGAAAAAGTTCCGAAGAGTGCATAAGAATATTGATTTG 179
Db 50697 ATGGATCCACTTCTCCATCGGAAGAGTTCATATTGTTGAACAAGATATTGATCAATT 50756
QY 180 GATTCTTTCACTCGTTGGTTCATGATGACCTCAAGAGTCCCTCAAAATATCAAAA 239
Db 50757 GATTGTTTGACGTTGTTTCATTTCTTATGAAGGTTTCCAAAGGTTCCCATAGCCTTA 50816
QY 240 GCGGAATCACAAATTTGAAATGTGATTGAATTCATTTTGTCTAATGCACAAAACAGGCA 299
Db 50817 GAGTTTTCGAATGAGATACGAATGCTACTCATCTCCAGAGAGGAATGAGAATG 50876
QY 300 TTCATAGCCTTTGTTTAAAGCAAAAACATTTCTTCCGATTTCATCCCA 349
Db 50877 TTTGCGCCTTCAATCAATTAATGAACCTTCTTTTGTCAATCATCACTCCA 50926

RESULT 15
AC144538/c 135983 bp DNA linear HTG 05-DEC-2003
LOCUS Medicago truncatula clone mth2-31121, WORKING DRAFT SEQUENCE, 3
DEFINITION
ordered pieces.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 135983)
Shaul, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Medicago truncatula BAC Clone mth2-31121
Unpublished
REFERENCE
2 (bases 1 to 135983)
Shaul, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (29-APR-2003) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
3 (bases 1 to 135983)
Shaul, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B.,
Cook, D., Kim, D. and Roe, B.A.
Direct Submission
Submitted (05-DEC-2003) Department of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT
On Dec 5, 2003 this sequence version replaced gi:38678558.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

```

* the accession number will be preserved.
* 1 104469: contig of 104469 bp in length
* 104470 104569: gap of unknown length
* 104570 113132: contig of 8563 bp in length
* 113133 113232: gap of unknown length
* 113233 135983: contig of 22751 bp in length.

Location/Qualifiers

1..135983
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/db_xref="taxon:3880"
/clone="mth2-31121"
/clone_lib="Medicago truncatula BAC library H2"

ORIGIN

Query Match 3.6%; Score 78.4; DB 2; Length 135983;
Best Local Similarity 53.4%; Pred. No. 1.2e-08;
Matches 186; Conservative 0; Mismatches 161; Indels 1; Gaps 1;

Qy 6 CCACTTTTAGGAATGGATCTTAAATTTTAGTTATAGTTCAAAGTTAGAAATCTTT 65
Db 114529 CCAAGATCTGGGAGACATCTCAAATTTTCTTACTCTTTCATGGATTGAAATCCTT 114470
Qy 66 ACCAAGAGCTTTGAGTCCATTGATGACATCGGTGAAACGGT-GTACATGTCTCCGATGGA 124
Db 114469 TTCAGAGAGCGTAACTCATTCATTAATGATAGTAAACCTTCCATACATTTCATCTATGGT 114410
Qy 125 CTCACCTGGTTTCATTCGGAAGAGTTGGAAGAGTGCATAGATATGATTTGGATTC 184
Db 114409 TTCGGTTTCTTTTCATTTCCGAATAGTTCAAACTTTGGAACACCAATGTCAATTTGTTTC 114350
Qy 185 TTTCACCTGGTTGGTGCCTTCATGAGTGCACCTCAGAGTCTCCAAATATCAAAGCCGA 244
Db 114349 CTTTACGTGACGTGGTGCCTTCATGATGATACCTTTAATGTATCCCAAATTTCTTGGCGTT 114290
Qy 245 ATCAAAATGAAATGTGATTAATGAAATTTTGTCTAATGCACAAACAGGCGATTCAT 304
Db 114289 TTTCATCTCTGACTCTATCATATTTCTCCCTGCTCAACGCCATGTTAGAAATAGTTT 114230
Qy 305 AGCCTTTGTGTTAAAGCAAAACATTTCTTCCGATTCATCCCATTC 352
Db 114229 AGCTTAAATTTAATAACATGCGCTTACCTTTCTTCTTGACCCAGTC 114182

Search completed: September 18, 2004, 23:51:57
Job time : 8138.53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 14:18:43 ; Search time 762.706 Seconds
(without alignments)
12030.986 Million cell updates/sec

Title: US-09-892-635A-45

Perfect score: 2160

Sequence: 1 ggatcccaacttttaggaat.....agctctctcattcaccatgg 2160

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2105.2	97.5	4924	2	AAX25614 Banana ri
2	73.4	3.4	809	2	Aav28690 Ripening
3	62.8	2.9	2417	2	Aav22222 SIRE-1 CD
4	54	2.5	110000	3	Continuation (3 of
5	53.4	2.5	110000	3	Continuation (4 of
6	52.6	2.4	515	2	Aac81475 HMG2 prom
7	51.4	2.4	2000	7	Ada71938 Rice gene
8	46.8	2.2	16439	6	Ab132886 Human imm
9	46.2	2.1	2000	7	Ada71938 Rice gene
10	43.4	2.0	742	2	Aav28672 Ripening
11	42.4	2.0	1186	2	Aax25608 Banana ri
12	41.6	1.9	5945	6	Ab132084 Human imm
13	41.6	1.9	7001	6	Abk33919 Human DNA
14	41.6	1.9	7001	7	Ada20394 Prostate
15	41.6	1.9	7001	7	Ada84201 Human ren
16	41.2	1.9	1722	7	AcF73552 Staphyloc
17	41.2	1.9	2881	2	Aav74565 Staphyloc
18	41.2	1.9	5879	6	Ab132269 Human imm
19	41.2	1.9	8197	6	Ab134515 Human met
20	41.2	1.9	8197	6	Ab170542 Chemical
21	41	1.9	6022	4	Aas46662 Tumour su
22	40.8	1.9	7351	6	Ab132028 Human imm
23	40.8	1.9	18967	4	Ab123560 Drosophil

24	40.6	1.9	726	6	ABQ68510	Listeria
25	40.6	1.9	1337	6	ABQ70519	Listeria
c 26	40.6	1.9	6216	6	ABK39932	Human che
c 27	40.6	1.9	6216	6	ABK39932	Human che
c 28	40.4	1.9	6161	6	ABL32623	Human imm
c 29	40.4	1.9	6237	6	ABL32623	Human imm
c 30	40.2	1.9	15399	6	ABL32925	Human imm
c 31	40	1.9	9483	6	ABL33515	Human imm
c 32	40	1.9	9483	6	ABL32377	Human imm
c 33	40	1.9	9483	6	ABL70516	Chemical
c 34	39.6	1.8	1722	7	ACA19762	Asa19762 Prokaryot
c 35	39.6	1.8	83391	6	ABQ67093	Abq67093 Human ang
c 36	39.4	1.8	9504	6	ABK28407	Abk28407 DNA trans
c 37	39.2	1.8	18679	6	ABN80300	Abn80300 Human che
c 38	39	1.8	5815	6	ABL33234	Human imm
c 39	39	1.8	5987	6	ABL33562	Human imm
c 40	39	1.8	5987	6	ABQ67101	Abq67101 Human ang
c 41	39	1.8	5987	9	ADB54307	Adb54307 Pretreate
c 42	39	1.8	5987	9	ADB54179	Adb54179 Pretreate
c 43	39	1.8	6089	4	AAS46579	Aas46579 Tumour su
c 44	39	1.8	6089	6	ABL33836	Ab133836 Human imm
c 45	39	1.8	12025	6	ABL33299	Ab133299 Human imm

ALIGNMENTS

RESULT 1

AAX25614

ID AAX25614 standard; DNA; 4924 BP.

XX AC AAX25614;

XX AC AAX25614;

DT 02-AUG-1999 (first entry)

XX XX Banana ripening fruit chitinase DNA.

DE DE Banana; fruit ripening; differential expression; fruit development;

XX KW transgenic plant; chitinase; ss.

XX OS Musa acuminata.

XX OS WO99156668-A2.

PN WO99156668-A2.

XX PD 01-APR-1999.

XX PF 23-SEP-1998; 98WO-US003343.

XX PF 25-SEP-1997; 97US-0060062P.

XX PR (BOYC-) BOYCE THOMPSON INST PLANT RES.

XX PA May G, Clendennen S;

XX PI WPI; 1999-244425/20.

XX DR P-PSDB; AAY05845, AAY05846, AAY05847.

XX XX New isolated banana DNA molecules.

XX PT Claim 5; Fig 17A-G; 143pp; English.

XX XX This is the nucleotide sequence of a chitinase DNA molecule. The

XX CC invention provides isolated DNA molecules which are differentially

XX CC expressed during banana fruit development, and the protein products of

XX CC these genes. The DNA is selected from a group comprising starch synthase,

XX CC chitinase, endochitinase, beta-1,3-glucanase, thaumatin-like protein,

XX CC ascorbate peroxidase, metallothionein, lectin and senescence-related

XX CC protein. The regulatory elements of the genes can be used to produce

XX CC chimeric genes for transformation of plants to provide controlled

XX CC expression of heterologous DNA during fruit development, or in response

XX CC to exogenous developmental signals, such as ethylene signals. The

XX CC heterologous protein, e.g. a therapeutic protein, can be isolated from

XX CC the fruit or consumed directly in the transformed fruit. The identity of

CC the present sequence is not given. The entire chitinase DNA sequence has
CC been decoded in all 3 reading frames to provide the amino acid sequences
CC given in AAY05845-47

XX
SQ Sequence 4924 BP; 1304 A; 1154 C; 1037 G; 1425 T; 0 U; 4 Other;

Query Match 97.5%; Score 2105.2; DB 2; Length 4924;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2148; Conservative 0; Mismatches 4; Indels 9; Gaps 3;

Qy	1	GGATCCCAACCTTTAGGAATCGATCTTAAATTTTAGTTATAAGTTCAAAAGTTAGAAA	60
Db	1	GGATCCCAACCTTTAGGAATCGATCTTAAATTTTAGTTATAAGTTCAAAAGTTAGAAA	60
Qy	61	TCCTTACCAAGAGCTTTGAGTCCATTTGATGACATCCGGTGAACCGGTGACATGCTCCGA	120
Db	61	TCCTTACCAAGAGCTTTGAGTCCATTTGATGACATCCGGTGAACCGGTGACATGCTCCGA	120
Qy	121	TGGACTCACTTGGTTTCATTCGGAAAAGTTCCGAAAGAGTGCATAGAAATATTGATTTGG	180
Db	121	TGGACTCACTTGGTTTCATTCGGAAAAGTTCCGAAAGAGTGCATAGAAATATTGATTTGG	180
Qy	181	ATTCCTTTCACCTCGGTGGTCTTCATGAGTGACCTCAAGAGTCTCTCCAAATATCAAAAG	240
Db	181	ATTCCTTTCACCTCGGTGGTCTTCATGAGTGACCTCAAGAGTCTCTCCAAATATCAAAAG	240
Qy	241	CCGAATCAAAATGGAATGATGAAATTCATTTTCTCTTAATGCAAAAACAGGGCAT	300
Db	241	CCGAATCAAAATGGAATGATGAAATTCATTTTCTCTTAATGCAAAAACAGGGCAT	300
Qy	301	TCATAGACCTTTGTTTAAAGCAAAAACATCTCTCCGATTCATCCCATTCGCTCATCG	360
Db	301	TCATAGACCTTTGTTTAAAGCAAAAACATCTCTCCGATTCATCCCATTCGCTCATCG	360
Qy	361	GAAGAGAAATTTTGAATCCATTTTCGCAATAGACCAAGGCTCGAAATCCATGCAATG	420
Db	361	GAAGAGAAATTTTGAATCCATTTTCGCAATAGACCAAGGCTCGAAATCCATGCAATG	416
Qy	421	GAAATGAGGAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACATCAATAAACATA	480
Db	417	GAAATGAGGAGATCCTCATATGAGTTTTCCAATACATGTAATTCGACATCAATAAACATA	476
Qy	481	GGTGATGTGTAATCAAAATGACCTCATGCSCTATCTCTTGGTATTAACCAAAATAT	540
Db	477	GGTGATGTGTAATGAAATGACCTCATGCSCTATCTCTTGGTATTAACCAAAATAT	536
Qy	541	GAGAGTGACCTTGTCTGATACCAATTTGATGATCAGAGTGACCACTAAGAGAGGGGG	600
Db	537	GAGAGTGACCTTGTCTGATACCAATTTGATGATCAGAGTGACCACTAAGAGAGGGGG	596
Qy	601	GAGTCAATGATGAGTGGATTAACCTTATAGTTTAAATGAAATGAAATCGTAAATACGAG	660
Db	597	GAGAGATGATGAGTGGATTAACCTTATAGTTTAAATGAAATGAAATCGTAAATACGAG	656
Qy	661	AAGATTTGGTTTAAATGATGAACTGAGTAGAGTGAACCAAAAGTTAAACAGTAGTGAAA	720
Db	657	AAGATTTGGTTTAAATGATGAACTGAGTAGAGTGAACCAAAAGTTAAACAGTAGTGAAA	716
Qy	721	TAACAATTTCCGGAAGTAAGAACTCACATATCAAGGAACATACCAATTTAAAGTGGTT	780
Db	717	TAACAATTTCCGGAAGTAAGAACTCACATATCAAGGAACATACCAATTTAAAGTGGTT	776
Qy	781	CGGTCAAAATGACCTACATCCACTGTGAGGCTTCTCGAGAGGCTCCCAACTCCAC	840
Db	777	CGGTCAAAATGACCTACATCCACTGTGAGGCTTCTCGAGAGGCTCCCAACTCCAC	836
Qy	841	TAGCAAAATCACCTTTGAAGGGGAGGACAAATACCTCTCTTACNACCTTTTCAATGGTTC	900
Db	837	TAGCAAAATCACCTTTGAAGGGGAGGACAAATACCTCTCTTACNACCTTTTCAATGGTTC	896
Qy	901	ATACCTTTCAAAATTTTCAACGAGAAAGAGAGTGAAATGGAAGCAATTTGAAACAA	960
Db	897	ATACCTTTCAAAATTTTCAACGAGAAAGAGAGTGAAATGGAAGCAATTTGAAACAA	956

Qy	961	GACTTGCTAAAGACTTTTGCTAAGGCTTTTCTCAATCTATCTCTCAAAAGTTGTA	1020
Db	957	GACTTGCTAAAGACTTTTGCTAAGGCTTTTCTCAATCTATCTCTCAAAAGTTGTA	1016
Qy	1021	TTCTCTGCTGAGAAATTCAGGGGTATTATAGACCCCAAGAGATTTAAATTTGGGCTCCA	1080
Db	1017	TTCTCTGCTGAGAAATTCAGGGGTATTATAGACCCCAAGAGATTTAAATTTGGGCTCCA	1076
Qy	1081	AAATTCGAAATGCTCTTGGGTTCCCGAGTTGCGGTGCGACCGCTGTCAGTGTGACAT	1140
Db	1077	AAATTCGAAATGCTCTTGGGTTCCCGAGTTGCGGTGCGACCGCTGTCAGTGTGACAT	1136
Qy	1141	CTGACAGTGTACTAGCGGTGCCACCGCGAACCTCTCGGTGTTGGCGGTGCCACCGC	1200
Db	1137	CTGACAGTGTACTAGCGGTGCCACCGCGAACCTCTCGGTGTTGGCGGTGCCACCGC	1196
Qy	1201	CTAGACTTTTTCAGCTCACTGGTTGGATTTCAAACTTGACCCCAACAGTCCGAACCTCG	1260
Db	1197	CTAGACTTTTTCAGCTCACTGGTTGGATTTCAAACTTGACCCCAACAGTCCGAACCTCG	1256
Qy	1261	GTCCAAATGACCCGTAACCGGATTTATAGGATTAACCTTAACTTAACCTTAATATATG	1320
Db	1257	GTCCAAATGACCCGTAACCGGATTTATAGGATTAACCTTAACTTAACCTTAATATATG	1316
Qy	1321	CAAACTAGCGAACTGAAATATAGTCTTAAGCAAGTTTAAACCGCAACAGTCCGAGTCT	1380
Db	1317	CAAACTAGCGAACTGAAATATAGTCTTAAGCAAGTTTAAACCGCAACAGTCCGAGTCT	1376
Qy	1381	TCCTCCGCGCATCTTTCGGCAGACTTCTGATATACCTTTGGATTTCTCTAGCGGACTCC	1440
Db	1377	TCCTCCGCGCATCTTTCGGCAGACTTCTGATATACCTTTGGATTTCTCTAGCGGACTCC	1436
Qy	1441	TAGTAGGTTCCCGATCTTTCGGCGAGTTTACGAGTAGCCGAACTTCTCGGTGATCTCC	1500
Db	1437	TAGTAGGTTCCCGATCTTTCGGCGAGTTTACGAGTAGCCGAACTTCTCGGTGATCTCC	1496
Qy	1501	GCAACCGCGCATGATCTTTCGGCAGACTTTCGAAACTTTTCGAACTTTCGAACTTTC	1560
Db	1497	GCAACCGCGCATGATCTTTCGGCAGACTTTCGAAACTTTTCGAACTTTCGAACTTTC	1556
Qy	1561	TCTCGGTTGGTTCCGACAGCATCTTAAACGAACTTTCGACTCTTGAATTCGAACTCGAA	1620
Db	1557	TCTCGGTTGGTTCCGACAGCATCTTAAACGAACTTTCGAACTTTCGAACTCGAA	1616
Qy	1621	CTTGACTCCGTTAGCTTCTTTATTTTTCAGGCTATCATAGTAACTTCTACATACATTA	1680
Db	1617	CTTGACTCCGTTAGCTTCTTTATTTTTCAGGCTATCATAGTAACTTCTACATACATTA	1676
Qy	1681	ACTCAATAATATGATGATTAATTAACCCATCAATGATTTTCATCATCAAAATTCGAC	1740
Db	1677	ACTCAATAATATGATGATTAATTAACCCATCAATGATTTTCATCATCAAAATTCGAC	1736
Qy	1741	ATTCAACAAACATCCGTTACTCAATAACCCATCAGGCTATAGTACGTAATCTACTGT	1800
Db	1737	ATTCAACAAACATCCGTTACTCAATAACCCATCAGGCTATAGTACGTAATCTACTGT	1796
Qy	1801	GATCGTTAGCTGAAGTTAGCGAGTCAATCGGTTCGTTCACTTATTTGGCGGAACAG	1860
Db	1797	GATCGTTAGCTGAAGTTAGCGAGTCAATCGGTTCGTTCACTTATTTGGCGGAACAG	1856
Qy	1861	TATCCCTTATCCAAATCCAGTCTTCTCAACTCTTCTAGCTTACCGCTCTCTTTTATTT	1920
Db	1857	TATCCCTTATCCAAATCCAGTCTTCTCAACTCTTCTAGCTTACCGCTCTCTTTTATTT	1916
Qy	1921	ACTTTTGAAGAAATTCAAATCAAAACAGATCAAAATAACAGGTGAGACATGTGACAT	1980
Db	1917	ACTTTTGAAGAAATTCAAATCAAAACAGATCAAAATAACAGGTGAGACATGTGACAT	1976
Qy	1981	GCTAGTCTCTGAAAGCAATTAATTCGCGCATCCAGACGCTGCTCAGCTTCTATCACCAC	2040
Db	1977	GCTAGTCTCTGAAAGCAATTAATTCGCGCATCCAGACGCTGCTCAGCTTCTATCACCAC	2036

135 TTCAATCGGAAAGTTCGAAAGAGTGCATAGAAATATTGATTTGGATCTTTCACTCGG 194
 122 TTCAATTTGAAGTTCATATCGTGAAGTGAGCAATCTTTGGATCTTTGACTGT 181
 195 TTGTGCTTCATGAGTGAGTCAAGAGTCCCTCAAAATATCAAAAGCCGAATCAAAAT 254
 182 TCAGTTTCTTCATGTGAGTCAAGCAATCCAGATTTCTTTAGCAGACTACAGGCT 241
 255 GAAATGATGATTAATTCATTTTGTCTTAATGCACAAACAGGCAATTCATAGCCTTTGTG 314
 242 GACACTATGTACTCATCAGGTCCTATCCACAGACCAATAGAGTTTACCTTTTGAAG 301
 315 TTTAAAGCAAAACATTTCTTCGATTCATCCCATTCGCTCATCGAAGAGAA 369
 302 CCCTTTCTATCTTTTCTCTGAGCAATCATATTTCTGCTTGGCTTTGGAA 356

RESULT 7
 ADA71938
 ID ADA71938 standard; DNA; 2000 BP.
 AC ADA71938;
 XX 20-NOV-2003 (first entry)
 DT Rice gene, SEQ ID 5263.
 DE
 DE Plant; bacterial infection; fungal infection; viral infection; rice;
 KW gene; ds.
 KW
 XX Oryza sativa.
 XX
 XX WO2003000898-A1.
 XX
 XX 03-JAN-2003.
 XX
 XX 22-JUN-2001; 2001WO-IB001105.
 XX
 XX 22-JUN-2001; 2001WO-IB001105.
 XX
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.
 XX
 XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 PS
 PS Claim 27; SEQ ID NO 5263; 899pp; English.
 XX
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX
 XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
 SQ

Query Match 2.4%; Score 51.4; DB 7; Length 2000;
 Best Local Similarity 8.5%; Pred.No. 0.0012;
 Matches 60; Conservative 325; Mismatches 322; Indels 1; Gaps 1;
 343 CATCCCATTCGCTCATCGGAAGAGAAATTTTGAATCCATTTTCGACAAATAGACAAA 402
 :| : : :| : : : : : : : : : : : : : : : :| : : : :

239 YMTTISWACSSYTWCRSKRRSMWKMWRKMSRSYGVWISWSYMMWCTAYKKSYSRW 298
 403 GCTCGAAATCCATGATGGAATGAGGAAGATCCTCATATGAGTTTTCATAATCATGTAA 462
 299 CYMYRGGGWRGATRYWGRGYMSRMAMMYKMYVRYGKMGKRWGAGBMMW-RSMCRWSK 357
 463 TTCGACTCATTAACATAGGTGATGTAATGAATGACCCCTCATCGCTATCTCTCTT 522
 358 ACYMRMRMRMRMTTTRRRWAKKSSRTSRKKRKWKMRKRYKMRGYSRMSRCKEARWKR 417
 523 GGGTATTAAACCAAAATATGAGAGTGAGCCTTGCTCATACCAATTTGTTAGGATCAGAGT 582
 418 CRSGRAWKMGCRGCMTCRMKSYGMRWKSWSKRYKMSRMYRWKRRKCKSRITTMGKT 477
 583 GGCACATAAGAGAGGGGGGAGTGAAATAGTGCAGTGGATTAATAAGTTTATAAAAA 642
 478 RGMWMTGRCRYKRSKMRKCRRRRWGRMYRWKRYMSARYTWRYCARKKYSYSA 537
 643 TGAATTCGTAATACGAGAAGATTTGTTTTTAATAGTAACCTTGAGTAGAGAAACCAA 702
 538 RKARWYRGKGYWAGWMMKRYKRYMYKMMWYKRYKSKSWYKMSYASCMKSARKA 597
 703 AGTTAACAGTAGTAAATAACAAATTTGGGAAAGTAAGAACTCACACATTCAGGGAACA 762
 598 GAKMCKRSKMSAMSKMSRSRCKRCASKRSSAKRYAMMGMTSGSRMSRWKSYTCVWRK 657
 763 TACCAATTTAAAGTGTTCGGTCAAAATAGACCTACATCCACTTGTGAAGCCTTTCTCGAA 822
 658 WGSMTKSTCTWYMYNSYTYAKYGSYWRVYRWCMYWRMYRYRYRYMYTYMAWYTSSTR 717
 823 GAGGCTCCCAACTTCCACTAGCAAACTCTTCAAGGGGAGGACAAATACCTCTCTTAC 882
 718 MAMTGMKYSGRYWTSWYKCYKSKYRSMYMSWMAKTWKMRYRYATRMWMMWYRYSM 777
 883 NACCTTTTACAATGGTTTCATCTTACAAATTTTCAACAGAAAGAGGAGGTGAACAT 942
 778 KWTWCTMGYWWYWRITYKMYMYKCTKYWYSATYWTGTWAAWMAKTMRMGMT 837
 943 GCAAGCAATGAAACAAAGACTTGTCTAAGACTTTTCTAAGGCTTTTCTCAATCTAT 1002
 838 GAKTRGRARKARYWKKWATWCAFKRWTKGKGAKWATWMAKAWRYKYSWMAWYKYK 897
 1003 TGTCTTCTCAAAAGTTGTTCTCTCTGCTGAGAATGAGGGGTATTATA 1050
 898 TRTRYKTCWKKARWGSWAYRWYRWMMKXGSAKMMWMMWKGGRWGTWYKYYW 945

RESULT 8
 ABL32886/C
 ID ABL32886 standard; DNA; 16439 BP.
 XX
 AC ABL32886;
 XX
 XX 26-MAR-2002 (first entry)
 DT Human immune system associated gene SEQ ID NO: 859.
 DE
 DE Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 KW ds.
 XX
 XX Homo sapiens.
 OS
 XX WO200200928-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX 02-JUL-2001; 2001WO-EP007537.
 PF

XX 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX (EPiG-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K;
 PI WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful for
 XX diagnosis and treatment of diseases associated with abnormal cytosine
 XX methylation.
 XX Claim 1; SEQ ID NO 859; 32pp + Sequence Listing; German.
 XX The present invention provides a number of human immune system associated
 XX genes which are modified by the methylation of cytosines. The sequences
 XX can be used in the diagnosis and treatment of immune system disorders,
 XX including eye diseases such as retinopathy, neovascular glaucoma and
 XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 XX diseases. The present sequence is a gene of the invention
 XX Sequence 16439 BP; 4389 A; 191 C; 3813 G; 8046 T; 0 U; 0 Other;
 SQ Query Match 2.2%; Score 46.8; DB 6; Length 16439;
 Best Local Similarity 55.6%; Pred. No. 0.074;
 Matches 90; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 639 AAAATGAATTCGTAATACGAGAGATTTCGTTTAAATAGTAAGTGGATGAGTGAAC 698
 DB 2835 AAAACGAAATATAAATCCAAAAAACCATTAAATTAATCTCCATTCAATAAAAA 2776
 QY 699 CAAAAGTTTACAGTAGTGTAAATAACAATTTCCGGGAAAGTAAAGAACTCACACATTCAAGG 758
 DB 2775 CACCAATTAATTAATAAAAAAACAATTTATTTCAATAAAAAATAAAAAATA 2716
 QY 759 AACATACCAATTTAAAGTGGTTCGGTCAAAATGACCTACATC 800
 DB 2715 AATATCCCTAATAAAAAAATAAATACTATAATCCCTACCTC 2674
 RESULT 9
 ID ADA71938 standard; DNA; 2000 BP.
 XX ADA71938;
 AC ADA71938;
 XX 20-NOV-2003 (first entry)
 XX Rice gene, SEQ ID 5263.
 XX Plant; bacterial infection; fungal infection; viral infection; rice;
 XX gene; ds.
 XX Oryza sativa.
 XX WO2003000898-A1.
 XX 03-JAN-2003.
 XX 22-JUN-2001; 2001WO-IB001105.
 XX 22-JUN-2001; 2001WO-IB001105.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
 PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to
 PT pathogenic infection for conferring resistance or tolerance to a plant to
 PT bacterial, fungal or viral infection by determining or detecting plant
 PT gene expression.
 XX Claim 27; SEQ ID NO 5263; 899pp; English.
 XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.
 XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
 SQ Query Match 2.1%; Score 46.2; DB 7; Length 2000;
 Best Local Similarity 9.9%; Pred. No. 0.037;
 Matches 85; Conservative 381; Mismatches 385; Indels 7; Gaps 3;
 QY 33 TTATAGTTTAAAGTTCAAAGTTAGAAAAATCTTTACCAAGAGCTTTGAGTCATTGATGAC 92
 DB 1018 TKWRTYTRMSWTYAMWKYKYKTYMTAYSTWYKWAYKRAYAWSRSRKTWCTGGKEM 959
 QY 93 ATCCGTGAAACGGTGTACATGTCCTCGATGGACTCACTGCTGTTTCATTCGGAAAAAGTTCG 152
 DB 958 ATYCGTKMAAGRWWRWMAWCYCMWKMKMTSCMWKYWRTWSCWTWNGAMRYAYY 899
 QY 153 AAAGAGTGCATAAGAAATATTCATTTTGGATTCTTTCACTCGG----TTGGTCCTTCATG 208
 DB 898 AMRRRTYTKWSWRMYTMTKWAATWMTCMMAKMYMATGWATMMWRMYTMYCYAMT 839
 QY 209 AGTGACCTCAAGAGTCTCCAAATATCAAAGCGGAATCACAATTGAAATGTGATGAA 268
 DB 838 CAKCKYKAMTKWTTWACAWRAISWRWRWAGRWKRYKMKRAYWWRWRCWAGWARW 779
 QY 269 TTCATTTTGTCTAAATGCACAAAACAGCGCAATCATAGCCTTTGTTTAAAGCAAAAC 328
 DB 778 MKSYRWKKKYATRYWKWAMTWNWSRRWRYWWSGMRWRSWRYCSRMCCKT 719
 QY 329 ATTCTTCTCCGATTCATCCCATTCGCTCATCGGAGAGAAATTTTGAATCCATTTC 388
 DB 718 KYASSARWTKRAKESYRYRRRWYWRKGTWYRYRYWRSRMTARMRSRRKRWAGASMKSC 659
 QY 389 GACAATAGACCAAGCTCGAAATCCATGCAATGAAATGAGGAAGATCCATATGAGTTT 448
 DB 658 WMY--WRGASRWYSKYSCSAKCKKTRYMTSSYSGTSGMYGMYSSYKMSWTSKMSYWGK 601
 QY 449 TCCAAATACATGTAATTCGACTCATTAACAATAGTGGATGTGTAATGAATGACCTCAT 508
 DB 600 MTCMTYTSKMGSTRRSKMGWSGMSRMYWKKWKKWRYKMYWKKCTWRRCMCYRWGY 541
 QY 509 GCSCTATCTCTTTGGGTATTAAACCAATATGAGAGTGAGCCTTGCTGTGATACCAATT 568
 DB 540 TMYTTSRSMYTGTRYKARYTSKRRYKRYKRYCWYYYGYMYKCSYMYRYGYCKACKK 481
 QY 569 GTTAGGATCAGATGGCACTAAGAGAGGGGGGAGTGAATTAGTGAGTGAATTAATAACT 628
 DB 480 CCYAMCWKAAYSGMMYWKYKSKWMSRSTKYMWSMYKCRSMKY-GAKCYGCKMWTY 422
 QY 629 TATAAGTTTAAAAATGAATTCGTAATACGAGAGATTTTCGTTTAAATAGTAACCTGAGT 688
 DB 421 CSYGYMKWYTYNGSYKYSRCYKMYMYKGMNMYMYYSMMTYYTYYKMYKYWKY 362
 QY 689 AGATGAAACCAAAAGTTAAACAGTAGTGTAAATAACAATTTTCGGGAAAGTGAAGAACTCA 748
 DB 361 RRGTSWYGKSKYKKYCTWCMYCMRCYRWKMKMKRKKTKYKRCYCHRYATCYWCCYR 302

one or more data sets. The genomic DNA is chemically treated by using a bisulphite, hydrogen sulphite or disulphite. The amplification step amplifies DNA which is of particular interest in astrocytoma or brain tissue, based on the specific genomic methylation status of brain tissues, as opposed to background DNA. The amplification carries a fluorescent label or radionuclide. Optionally, the labels of the amplicates are detachable molecule fragments having a typical mass which are detected in a mass spectrometer. The fragments of chemically pre-treated genomic DNA to be amplified, have a single positive or negative charge for a better detectability in the mass spectrometer. Preferably, the amplicates or fragments of the amplicates are detected by matrix assisted laser desorption/ionization mass spectrometry (MALDI) or using electron spray mass spectrometry (ESI). The present sequence is one of the chemically pre-treated reference DNA samples of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 7001 BP; 2309 A; 40 C; 1249 G; 3403 T; 0 U; 0 Other;

Query Match 1.9%; Score 41.6; DB 6; Length 7001;
Best Local Similarity 48.0%; Pred. No. 1.5;
Matches 119; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 229 AATATCAAAAGCGGATCAAAATGGAATGTAATTCATTTTGTCTAATGCAC 288
DB 667 AAAATTCAAAATAAATACTTATCCCTAAATACAAAATAAATAATTCATTCAATATC 608
QY 289 AAAACAGGGGATTCATAGCCTTTGTTTAAAGCAAAACATTTCTCCGATTCATCCC 348
DB 607 ATTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 548
QY 349 ATTGCTCATCGGAGAGAGAAAATTTTGAATCCATTTTCGACATAGACCAAGCTCGA 408
DB 547 ATCAATAAATAACATTTTAAATATTAATATAGATATATTAATTAACATAAAACTT 488
QY 409 AATCCATGATGAATGAGAGAGATCCTCATATGAGTTTCCAAATACATGTAATTCGAC 468
DB 487 AAAACACAAACCCCTAATAAACCAATCCTATAATAATTTATTACACATATATCTTAT 428
QY 469 TCATTAAA 476
DB 427 AAAATAAA 420

RESULT 14
ADA20394/c

ID ADA20394 standard; DNA; 7001 BP.

AC ADA20394;

DT 20-NOV-2003 (first entry)

DE Prostate tumour related genomic DNA sample #30.

XX cytostatic; gene therapy; genetic marker; epigenetic parameter;
XX classification; differentiation; diagnosis; prostate tumour;
XX prostate cancer; cytosine methylation; uracil;
XX single nucleotide polymorphism; SNP; prostate carcinoma; ss.

OS Homo sapiens.

XX WO2002103042-A2.

XX 27-DEC-2002.

XX 14-JUN-2002; 2002WO-EP006605.

XX 14-JUN-2001; 2001DE-01028508.

XX (EPIC-) EPIGENOMICS AG.

PI Distler J, Model F, Adorjan P;

XX

DR WPT; 2003-167536/16.

XX Determining genetic and/or epigenetic parameters, useful for the
PT classification, differentiation and/or diagnosis of prostate tumors or a
PT predilection to prostate cancer, comprises analyzing cytosine
PT methylation.

XX Claim 28; Page 262-266; 376pp; English.

XX The invention relates to a method of determining genetic and/or
CC epigenetic parameters for the classification, differentiation and/or
CC diagnosis of prostate tumors or the predilection to prostate cancer,
CC by analysing cytosine methylation in a sample of genomic DNA. The method
CC comprises chemically treating unmethylated cytosine bases at the 5-
CC position to uracil or another base, which is dissimilar to cytosine in
CC terms of hybridization behaviour; followed by amplifying at least one
CC fragment of the chemically pre-treated genomic DNA using sets of primer
CC oligonucleotides and a polymerase. The oligomers or probes derived from
CC them are useful for detecting the methylation state of all CpG
CC dinucleotides and/or single nucleotide polymorphisms (SNPs) in a
CC chemically pre-treated genomic DNA. They are all useful for treating
CC prostate carcinoma. This sequence represents a fragment of genomic DNA
CC used in the method of the invention.

XX SQ Sequence 7001 BP; 2309 A; 40 C; 1249 G; 3403 T; 0 U; 0 Other;

Query Match 1.9%; Score 41.6; DB 7; Length 7001;
Best Local Similarity 48.0%; Pred. No. 1.5;
Matches 119; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 229 AATATCAAAAGCGGATCAAAATGGAATGTAATTCATTTTGTCTAATGCAC 288
DB 667 AAAATTCAAAATAAATACTTATCCCTAAATACAAAATAAATAATTCATTCAATATC 608
QY 289 AAAACAGGGGATTCATAGCCTTTGTTTAAAGCAAAACATTTCTCCGATTCATCCC 348
DB 607 ATTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 548
QY 349 ATTGCTCATCGGAGAGAGAAAATTTTGAATCCATTTTCGACATAGACCAAGCTCGA 408
DB 547 ATCAATAAATAACATTTTAAATATTAATATAGATATATTAATTAACATAAAACTT 488
QY 409 AATCCATGATGAATGAGAGAGATCCTCATATGAGTTTCCAAATACATGTAATTCGAC 468
DB 487 AAAACACAAACCCCTAATAAACCAATCCTATAATAATTTATTACACATATATCTTAT 428
QY 469 TCATTAAA 476
DB 427 AAAATAAA 420

RESULT 15
ADA84201/c

ID ADA84201 standard; DNA; 7001 BP.

XX AC ADA84201;

XX DT 20-NOV-2003 (first entry)

XX Human renal/prostate carcinoma associated DNA SEQ ID NO:59.

XX ds; renal cancer; prostate cancer; cytosine methylation;
XX single nucleotide polymorphism; histological; cytological.

XX OS Homo sapiens.

XX WO2002103041-A2.

XX 27-DEC-2002.

XX 14-JUN-2002; 2002WO-EP006603.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 11:51:07 ; Search time 152.641 Seconds
(without alignments)
7853.010 Million cell updates/sec

Title: US-09-892-635A-45

Perfect score: 2160

Sequence: 1 ggatccaaacttttaggaat.....agctctctcattcaccatgg 2160

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	62.8	2.9	2417	US-09-254-776B-3	Sequence 3, Appli
C 2	52.6	2.4	515	US-08-282-581-5	Sequence 5, Appli
C 3	52.6	2.4	515	US-08-550-544-5	Sequence 5, Appli
C 4	48.4	2.2	7218	US-08-232-463-14	Sequence 14, Appl
C 5	44.8	2.1	7218	US-08-232-463-14	Sequence 14, Appl
C 6	42.8	2.0	832	US-09-621-976-2813	Sequence 2813, Ap
C 7	41.2	1.9	2881	US-08-956-171E-254	Sequence 254, App
C 8	37.6	1.7	5335	US-10-204-708-17	Sequence 17, Appl
C 9	37	1.7	6317	US-10-204-708-11	Sequence 11, Appl
C 10	36.6	1.7	399	US-09-621-976-8976	Sequence 8976, Ap
C 11	36.6	1.7	2871	US-09-023-655-768	Sequence 768, App
C 12	36.2	1.7	1797	US-09-134-000C-1552	Sequence 1552, Ap
C 13	36.2	1.7	6113	US-10-204-708-13	Sequence 13, Appl
C 14	35.8	1.7	6124	US-08-213-419B-3	Sequence 3, Appli
C 15	35.8	1.7	21234	US-09-810-671-3	Sequence 3, Appli
C 16	35.8	1.7	21234	US-10-109-854-3	Sequence 854-3
C 17	35.6	1.6	1704	US-10-134-001C-925	Sequence 925, App
C 18	35.6	1.6	168575	US-09-426-290-1	Sequence 1, Appli
C 19	35.4	1.6	4185	US-09-417-485D-7	Sequence 7, Appli
C 20	35.4	1.6	10640	US-09-417-485D-5	Sequence 5, Appli
C 21	35.2	1.6	11050	US-10-204-708-86	Sequence 86, Appl
C 22	35	1.6	6656	US-10-204-708-76	Sequence 76, Appl
C 23	35	1.6	10467	US-10-204-708-2	Sequence 2, Appli
C 24	35	1.6	1664976	US-08-916-421B-1	Sequence 1, Appli
C 25	34.8	1.6	3440	US-08-471-791-27	Sequence 27, Appl
C 26	34.8	1.6	3440	PCT-US91-01746-27	Sequence 27, Appl
C 27	34.8	1.6	5562	US-10-204-708-64	Sequence 64, Appl

C 28	34.6	1.6	627	US-09-495-050A-74	Sequence 74, Appl
C 29	34.6	1.6	6020	US-10-204-708-8	Sequence 8, Appli
C 30	34.4	1.6	832	US-09-621-976-2813	Sequence 2813, Ap
C 31	34.4	1.6	887	US-09-602-848-3	Sequence 3, Appli
C 32	34.4	1.6	99500	US-09-798-096-10	Sequence 10, Appl
C 33	34.2	1.6	337	US-09-985-799-79	Sequence 79, Appl
C 34	34.2	1.6	337	US-08-594-031-79	Sequence 79, Appl
C 35	34.2	1.6	2427	US-09-134-001C-269	Sequence 268, App
C 36	34.2	1.6	4673	US-07-638-431-1	Sequence 1, Appli
C 37	34.2	1.6	4673	PCT-US92-00018-1	Sequence 1, Appli
C 38	34.2	1.6	5152	US-10-204-708-74	Sequence 74, Appl
C 39	34.2	1.6	580073	US-08-545-528B-1	Sequence 1, Appli
C 40	34.2	1.6	640681	US-09-790-988-1	Sequence 1, Appli
C 41	34	1.6	505	US-09-621-976-15639	Sequence 15639, A
C 42	34	1.6	576	US-09-621-976-16902	Sequence 16902, A
C 43	34	1.6	1991	US-08-235-836C-69	Sequence 69, Appl
C 44	34	1.6	1991	US-08-235-836C-77	Sequence 77, Appl
C 45	34	1.6	5521	US-08-975-762-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1
US-09-254-776B-3/c
; Sequence 3, Application US/09254776B
; Patent No. 6559359
; GENERAL INFORMATION:
; APPLICANT: Laten, Howard
; TITLE OF INVENTION: PLANT RETROVIRAL POLYNUCLEOTIDES AND METHODS FOR USE THEREOF
; FILE REFERENCE: 27013/33479A
; CURRENT APPLICATION NUMBER: US/09/254,776B
; CURRENT FILING DATE: 1999-03-09
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 2417
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: SIRE-1 cDNA clone
US-09-254-776B-3

Query Match	2.9%	Score 62.8;	DB 4;	Length 2417;
Best Local Similarity	50.6%;	Pred. No. 9.7e-09;		
Matches	177;	Conservative	0;	Mismatches 172; Indels 1; Gaps 1;
QY	4	TCCCACTTTTAGGAATGGATCTTAAATTTTATAGTTTAAAGTTCAAGTTAGAAAATCT	63	
DB	786	TCAATCTCTTAGGCAAGGATCTGAGGATCTTCTCACAGCTTTTCATCAGTCATTCT	727	
QY	64	TTACCAAGAGCTTTGAGTCCATTGATGACATC-CGTGAACCGGTGTACATGTCCTCGATG	122	
DB	726	TCTCCCAAGGAGTCAAGCATTCGCAATTCAGAAATGTTTCATGTGAAGTCATGAATA	667	
QY	123	GACTCACTGGTTTTCATTCGGAAAAGTTGCGAAGAGTGCATGAATATTTGATTTGGAT	182	
DB	666	CACCTCTCTCTCTTCATCTTCAGATTTTCGAATTTTGTGGCCAAATAGTTGCAATCTGGAC	607	
QY	183	TCTTTCACTCGTTTGGTGGCTTCATGAGTGCCTCAAGATCTTCAATATCAAAAGCC	242	
DB	606	ATCTTCATCTTTGGAGGTTCCCTTCATGAGTGGTTTTCAGGATCTCCCATGCTCTTGGCC	547	
QY	243	GAATCACAATTAATGATGATTTGATTTTCTTAATGACACAAACAGGCAATTC	302	
DB	546	ACTGTGCATGTTTCATCAGTCTGAGATATTTCTTCACTCCATTTGATGAGCATTC	487	
QY	303	ATAGCTTTGTGTTTAAAGCAAAAACATTTCTCCGATTTTCATCCATTC	352	
DB	486	AAAGCTTTGGAGTTTCCCAAGTGCCTTCGTTCTTCTTTAGTCCAGTC	437	

Mon Sep 20 12:58:19 2004

```

RESULT 2
US-08-282-581-5
; Sequence 5, Application US/08282581
; Patent No. 5670349
; GENERAL INFORMATION:
; APPLICANT: Cramer, Carole L.
; APPLICANT: Weissenborn, Deborah L.
; TITLE OF INVENTION: HM2 PROMOTER EXPRESSION SYSTEM AND
; TITLE OF INVENTION: POST-HARVEST PRODUCTION OF GENE PRODUCTS
; TITLE OF INVENTION: IN PLANTS AND PLANT CELL CULTURES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,581
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18872
; REFERENCE/DOCKET NUMBER: 7956-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-282-581-5

Query Match 2.4%; Score 52.6; DB 1; Length 515;
Best Local Similarity 51.5%; Pred. No. 5,1e-06;
Matches 121; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Qy 135 TTCATTCGGAAGAAGTTCGAAGAGAGTGCAATAGATATTGATTTGGATCTTTCACCTGG 194
Db 122 TTCATTTTGAAGTTCTCATATCGTGAGGTGAGCATGTCAAATCTTGGATCTTTCGACTTGT 181
Qy 195 TTGGTGCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAGCCGAATCACAAATT 254
Db 182 TCAGTTCTTCAATGTTAGTCAAGCAATCCAGATTTCTTTAGCAGACTCAGGCT 241
Qy 255 GAAATGTGATTGAATTCATTTTGTCTAATGCACAAAACAGGGCAATCATAGCCTTTGTG 314
Db 242 GACACTCTATTGTACTCATCAGGTCCATCCACAGACCATAAAGAGTTTACCTTTGAGG 301
Qy 315 TTTAAGCAAAAACATTTCTCCGATTCATCCCATTCGCTCATCGGAAGAAA 369
Db 302 CCCTTTTCTATCTTTTCTCTGTCAGCATCATCATATTCTGCTTGGGCTTTGGAA 356

RESULT 3
US-08-550-544-5
; Sequence 5, Application US/08550544
; Patent No. 5689056
; GENERAL INFORMATION:
; APPLICANT: Cramer, Carole L.
; APPLICANT: Weissenborn, Deborah L.
; TITLE OF INVENTION: HM2 PROMOTER EXPRESSION SYSTEM

```


; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 768:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTUT01
; CLONE: 602872
US-09-023-655-768

Query Match 1.7%; Score 36.6; DB 4; Length 2871;
Best Local Similarity 45.5%; Pred. No. 1.2;
Matches 126; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
Qy 514 ATCTCTTGGGTATTAACCAATATGAGAGTGAGCCTTCTCTGATACCAATTGTTAG 573
Db 717 ATATTACAGGGTCAAGAAACAAAGCTGTCTCCAGTCATGTTTGGACAATAAAGCTTTG 658
Qy 574 GATCAGAGTGGCACTAAGAGAGGGGGAGTGAATTAGTCAGTGCATTAAACCTTAA 633
Db 657 GGGTCAGAGCGGAAGAGGAGGAAGAAAGGAAGAGAGGAGAAATACACTACTTT 598
Qy 634 GTTTAAAGTAAGTTCGTAATACGAGAGAGATTTCTGTTTTAATAGTAACTTGAGTAG 693
Db 597 CTGGAAACACATTTCGCTTAACTCCAAATAAAGGCTTTGCGGAGAAATGAAAGCCT 538
Qy 694 AAACCAAGTAACTAGTAGTGAATAACAAATTCGGGAAGTAAGAACTCACACATT 753
Db 537 ATAATCAGATTAGTGTGCAATAAACAACACAGCTGCACACAGACAATATCTACATCAT 478
Qy 754 CAAGCAACATACCAATTTAAAGTGGTTCGGTCAAAAT 790
Db 477 TAATANCCTTCNATTATAGATCTATCGTCACAAAT 441

RESULT 12
US-09-134-000C-1552

; Sequence 1552, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1552
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1552

Query Match 1.7%; Score 36.2; DB 4; Length 1797;
Best Local Similarity 49.2%; Pred. No. 1.2;
Matches 95; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
Qy 926 AAGAAGCAGGTGAACATCAAGCAATTGAAAAACAAGACTTGCTAAAGACTTTGCTAAGGC 985
Db 761 AAGAAGCAGACCAAGAAAGCTAATGAATTATGAAATGGTTCGCTAGCGAGTCTCTAAAA 820
Qy 986 TTTTCTTCTCAATCTATTGCTTCTCAAAAGTTGTTCTCTGCTGAGAAATGAGGGTAT 1045
Db 821 AATTGATAAAATCTATTCCATGTTAGAGTGTGAATTTGATTTCTTATAATGGCGAAGCCT 880
Qy 1046 TTATAGACCCCAAGAGGATTTAAATTTGGGTCCTCAAAATTTGGAATGCTCTGGGCTCCCG 1105
Db 881 TTATAATGATAAAATGATGAAATTTGTTACTTTATTAGAGAAAAACAATTATTACAG 940
Qy 1106 AGGTTGCGGTGC 1118
Db 941 AAAATCAAGGTGC 953

RESULT 13
US-10-204-708-13
; Sequence 13, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 13
; LENGTH: 6113
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-13

Query Match 1.7%; Score 36.2; DB 4; Length 6113;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 14:07:43 ; Search time 968.897 Seconds
(without alignments)
11252.468 Million cell updates/sec

Title: US-09-892-635A-45

Perfect score: 2160

Sequence:

1 ggatcccaacttttaggaat.....agactctcattcaccatgg 2160

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3327077 seqs, 2523723180 residues

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2158.6	99.9	2160	13	US-09-892-635A-45
2	2139	99.0	2156	13	US-09-892-635A-44
3	2110	97.7	4924	13	US-09-892-635A-27
4	122.2	5.7	78333	16	US-10-298-122-3
5	81.4	3.8	2760	13	US-10-424-599-108338
6	81.2	3.8	5895	16	US-10-298-122-2
7	73.4	3.4	809	9	US-09-966-881-48
8	62.8	2.9	2417	15	US-10-334-703-3
9	62.8	2.9	2417	16	US-10-396-122-3
10	62.8	2.9	2801	13	US-10-424-599-108650
11	62.8	2.9	9399	16	US-10-396-122-93
12	61.2	2.8	5371	17	US-10-437-963-9498
13	59.6	2.8	9072	16	US-10-396-122-87
14	59.4	2.8	5016	17	US-10-437-963-10782

C 15	59.4	2.8	5304	17	US-10-437-963-44132	Sequence 44132, A
C 16	57.8	2.7	3843	17	US-10-437-963-10117	Sequence 10117, A
C 17	57.8	2.7	4326	17	US-10-437-963-9179	Sequence 9179, Ap
C 18	57.8	2.7	4428	17	US-10-437-963-44137	Sequence 44137, A
C 19	57.8	2.7	4840	17	US-10-437-963-9223	Sequence 9223, Ap
C 20	57.8	2.7	10369	17	US-10-437-963-10605	Sequence 10605, A
C 21	57.6	2.7	4503	17	US-10-437-963-8782	Sequence 8782, Ap
C 22	56.6	2.6	7628	13	US-10-424-599-108640	Sequence 108640, Ap
C 23	56.2	2.6	4002	17	US-10-437-963-9586	Sequence 9586, Ap
C 24	56.2	2.6	4428	17	US-10-437-963-9621	Sequence 9621, Ap
C 25	56	2.6	3127	17	US-10-437-963-44134	Sequence 44134, A
C 26	55.6	2.6	9114	17	US-10-437-963-10728	Sequence 10728, A
C 27	55.6	2.6	2160	17	US-10-437-963-101688	Sequence 101688, A
C 28	54.8	2.5	846	17	US-10-437-963-96678	Sequence 96678, A
C 29	54.6	2.5	4884	17	US-10-437-963-9664	Sequence 9664, Ap
C 30	54.4	2.5	5370	17	US-10-437-963-9697	Sequence 9697, Ap
C 31	53.8	2.5	1731	17	US-10-437-963-9134	Sequence 9134, Ap
C 32	53.4	2.5	4382	17	US-10-437-963-44138	Sequence 44138, A
C 33	53.4	2.5	4716	17	US-10-437-963-10750	Sequence 10750, A
C 34	53.2	2.5	13744	17	US-10-437-963-11092	Sequence 11092, A
C 35	53	2.5	4125	17	US-10-437-963-9312	Sequence 9312, Ap
C 36	53	2.5	4620	17	US-10-437-963-93128	Sequence 10128, A
C 37	53	2.5	6036	17	US-10-437-963-44135	Sequence 44135, A
C 38	52.6	2.4	515	9	US-09-902-653-5	Sequence 5, Appli
C 39	52.6	2.4	515	13	US-10-394-494-5	Sequence 5, Appli
C 40	52.4	2.4	5856	17	US-10-437-963-23904	Sequence 23904, A
C 41	52	2.4	5208	17	US-10-437-963-34842	Sequence 34842, A
C 42	51.8	2.4	3810	17	US-10-437-963-24004	Sequence 24004, A
C 43	51.4	2.4	5092	17	US-10-437-963-10162	Sequence 10162, A
C 44	51.2	2.4	3231	17	US-10-437-963-10084	Sequence 10084, A
C 45	51.2	2.4	4420	17	US-10-437-963-9267	Sequence 9267, Ap

ALIGNMENTS

RESULT 1

US-09-892-635A-45

; Sequence 45, Application US/09892635A

; Publication No. US20030226175A1

; GENERAL INFORMATION:

; APPLICANT: May, Gregory D.

; APPLICANT: Clendennen, Stephanie K.

; APPLICANT: Mason, Hugh S.

; APPLICANT: Lim, Miguel A. Gomez

; APPLICANT: Arntzen, Charles J.

; TITLE OF INVENTION: DNA Regulatory Elements Associated with Fruit Development

; FILE REFERENCE: 031998-007

; CURRENT APPLICATION NUMBER: US/09/892,635A

; CURRENT FILING DATE: 2001-06-28

; PRIOR APPLICATION NUMBER: US 09/160,351

; PRIOR FILING DATE: 1998-09-25

; PRIOR APPLICATION NUMBER: US 60/060,062

; PRIOR FILING DATE: 1997-09-25

; NUMBER OF SEQ ID NOS: 45

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 45

; LENGTH: 2160

; TYPE: DNA

; ORGANISM: Musa acuminata

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 883

; OTHER INFORMATION: n = A,T,C or G

US-09-892-635A-45

Query Match

Best Local Similarity 99.9%; Score 2158.6; DB 13; Length 2160;

Matches 2160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 GGATCCCACTTTTAGGAATGGATCTTAAATTTTAGTTATAGTTCAAGTTAGAAAA 60

DB

1 GGATCCCACTTTTAGGAATGGATCTTAAATTTTAGTTATAGTTCAAGTTAGAAAA 60

Qy	61	TCTTTACCAAGAGCTTTGAGTCCATTGATGACATCGTGAAACGGGTACATGTCTCCGA	120
Db	61	TCCTTTACCAAGAGCTTTGAGTCCATTGATGACATCGTGAAACGGGTACATGTCTCCGA	120
Qy	121	TGGACTCACCTGGTGTTCATTTCGAAAAGTTCGAAAGAGTGCATAAGAAATTATTGATTTGG	180
Db	121	TGGACTCACCTGGTGTTCATTTCGAAAAGTTCGAAAGAGTGCATAAGAAATTATTGATTTGG	180
Qy	181	ATTCTTTCACTCGGTGGTGCTTCATGAGTGAACCTCAAGAGTCTCCAAAATATCAAAAG	240
Db	181	ATTCTTTCACTCGGTGGTGCTTCATGAGTGAACCTCAAGAGTCTCCAAAATATCAAAAG	240
Qy	241	CCGAATCACAAAATTGAAATGTGATGCAATTCATTTTGTCTAATGCACAAAAACAGGCAT	300
Db	241	CCGAATCACAAAATTGAAATGTGATGCAATTCATTTTGTCTAATGCACAAAAACAGGCAT	300
Qy	301	TCATAGCCTTTGTGTTTAAAGCAAAAAATCTTCTTCCGATTCAATCCCATTCGCTCATCG	360
Db	301	TCATAGCCTTTGTGTTTAAAGCAAAAAATCTTCTTCCGATTCAATCCCATTCGCTCATCG	360
Qy	361	GAGAGAAAATTTTGAATCCATTTTTCGACAATGACCAAGCTCGAAATCCATGCATG	420
Db	361	GAGAGAAAATTTTGAATCCATTTTTCGACAATGACCAAGCTCGAAATCCATGCATG	420
Qy	421	GAATGAGGAAGATCCCTCATATGAGTTTCCCAATACATGTAATTCGACTCATTAACATA	480
Db	421	GAATGAGGAAGATCCCTCATATGAGTTTCCCAATACATGTAATTCGACTCATTAACATA	480
Qy	481	GGTGGATGTGTAATGAAATGACCTCATGSCCTATCTCTCTGGGTATTAAACCAAAATAT	540
Db	481	GGTGGATGTGTAATGAAATGACCTCATGSCCTATCTCTCTGGGTATTAAACCAAAATAT	540
Qy	541	GAGAGTGAAGCTTCCTCTGATACCAATTTGTTAGATCAGAGTGGCACTAAGACAGGGGG	600
Db	541	GAGAGTGAAGCTTCCTCTGATACCAATTTGTTAGATCAGAGTGGCACTAAGACAGGGGG	600
Qy	601	GAGTGAATTTAGTGCAGTGGATTAAACTTATAAGTTTAAAAATGAAATCGTAAATACGAG	660
Db	601	GAGTGAATTTAGTGCAGTGGATTAAACTTATAAGTTTAAAAATGAAATCGTAAATACGAG	660
Qy	661	AAGATTTGCTTTTAAATAGTAATCTGAGTAGATGAAACCAAAAGTTTAACGATGCTGAAA	720
Db	661	AAGATTTGCTTTTAAATAGTAATCTGAGTAGATGAAACCAAAAGTTTAACGATGCTGAAA	720
Qy	721	TAACAATTTCCGGAAAGTAAACAATCAACATTTCAAGGAAATACCAATTTAAAGTGGTT	780
Db	721	TAACAATTTCCGGAAAGTAAACAATCAACATTTCAAGGAAATACCAATTTAAAGTGGTT	780
Qy	781	CGGTCAAATACCTTACATCCATTTGTGAAGCCTTCTTCGAAGAGGCTCCAACTTCCAC	840
Db	781	CGGTCAAATACCTTACATCCATTTGTGAAGCCTTCTTCGAAGAGGCTCCAACTTCCAC	840
Qy	841	TAGCAAAATCACTTTGAAGGGAAGCAAAATACCTCTTTACNACCTTTTACAAATGGTTC	900
Db	841	TAGCAAAATCACTTTGAAGGGAAGCAAAATACCTCTTTACNACCTTTTACAAATGGTTC	900
Qy	901	ATACTCTTACAAATTTTCAACGAGAAAGAGGGTGAACATCAAGCAATTCGAAAACAA	960
Db	901	ATACTCTTACAAATTTTCAACGAGAAAGAGGGTGAACATCAAGCAATTCGAAAACAA	960
Qy	961	GACTTGCTTAAAGACTTTTCTTAAAGCCTTTTCTCAATCTATTGCTTCTCAAAAGTTGTA	1020
Db	961	GACTTGCTTAAAGACTTTTCTTAAAGCCTTTTCTCAATCTATTGCTTCTCAAAAGTTGTA	1020
Qy	1021	TTCCTCTGCTGAGAAATTGAGGGGTAATTAAGACCCCAAGAGGATTTAAATTTGGGTCCA	1080
Db	1021	TTCCTCTGCTGAGAAATTGAGGGGTAATTAAGACCCCAAGAGGATTTAAATTTGGGTCCA	1080
Qy	1081	AAATTTCGAATGTCTTTGGGTTCCCGAGTTCGCGGTGCGACCGGCTGTGAGTTTGACA	1140
Db	1081	AAATTTCGAATGTCTTTGGGTTCCCGAGTTCGCGGTGCGACCGGCTGTGAGTTTGACA	1140

QY	1141	CTGGACAGTGTACTAGCGGTGCCACCGCGGACCTCTCGGGTGTTCGGCGGTGCCACCGC	1200
DB	1141	CTGGACAGTGTACTAGCGGTGCCACCGCGGACCTCTCGGGTGTTCGGCGGTGCCACCGC	1200
QY	1201	CTAGACTTTTTCAGCTCACGTGGTTGGATTCCAAACTTGACCCAAACAGGTCCGAACTCGG	1260
DB	1201	CTAGACTTTTTCAGCTCACGTGGTTGGATTCCAAACTTGACCCAAACAGGTCCGAACTCGG	1260
QY	1261	GTCCAATTGACCCGGTAACCGGATATTAGGATTAACCTTTAATCCTTAACCTTAATTTATNG	1320
DB	1261	GTCCAATTGACCCGGTAACCGGATATTAGGATTAACCTTTAATCCTTAACCTTAATTTATNG	1320
QY	1321	CAAACTAGCAACTGAAATATAGTCTTAAGCAAGTTTTTAAACGGCGAAACGTCGAGTCT	1380
DB	1321	CAAACTAGCAACTGAAATATAGTCTTAAGCAAGTTTTTAAACGGCGAAACGTCGAGTCT	1380
QY	1381	TCCTCCGCGACTTTTCGGCAGACTTCCTGATATACCTTTGGATTTCTTTAGCGGACTCC	1440
DB	1381	TCCTCCGCGACTTTTCGGCAGACTTCCTGATATACCTTTGGATTTCTTTAGCGGACTCC	1440
QY	1441	TAGTAGGGTCCGACTTTGTGGCGAGTTTAGCGAGTAGCGGAACTTCTCGGTGATCTCC	1500
DB	1441	TAGTAGGGTCCGACTTTGTGGCGAGTTTAGCGAGTAGCGGAACTTCTCGGTGATCTCC	1500
QY	1501	GCAAAACCGCGATGATCTTTCGGCAGACTTTCGAAACCTTCGAAAGTCCCGGATTTCT	1560
DB	1501	GCAAAACCGCGATGATCTTTCGGCAGACTTTCGAAACCTTCGAAAGTCCCGGATTTCT	1560
QY	1561	TCTCGGTGGTTTCGGACAGACTCTTAACGAAACTTCGGACTCCTTGAATGTCATCGAA	1620
DB	1561	TCTCGGTGGTTTCGGACAGACTCTTAACGAAACTTCGGACTCCTTGAATGTCATCGAA	1620
QY	1621	CTTGACTCCGGTAGGCTTGCTTTATATTTTCAGGCTATCATAGTTAACTCTCAATACTTA	1680
DB	1621	CTTGACTCCGGTAGGCTTGCTTTATATTTTCAGGCTATCATAGTTAACTCTCAATACTTA	1680
QY	1681	ACTCAATTAATAGATTAGATTAACTTAACCAATCAATTTTCATCATCAAAATTCGAC	1740
DB	1681	ACTCAATTAATAGATTAGATTAACTTAACCAATCAATTTTCATCATCAAAATTCGAC	1740
QY	1741	ATTCAACAAACATCCGTACTCAATTAACCCATCAGGCTATAGTTACGTGACTATCTAGT	1800
DB	1741	ATTCAACAAACATCCGTACTCAATTAACCCATCAGGCTATAGTTACGTGACTATCTAGT	1800
QY	1801	GATCGGTACGTGAAGTTAGCGAGTCATGATCCAGGTCGTGCATTAATTCGCCGAAACAG	1860
DB	1801	GATCGGTACGTGAAGTTAGCGAGTCATGATCCAGGTCGTGCATTAATTCGCCGAAACAG	1860
QY	1861	TATCCCTTATCCAAATCCAGTCTTCTCAACTCTTCTAGCCTACCGCTCTCTTTTTTATT	1920
DB	1861	TATCCCTTATCCAAATCCAGTCTTCTCAACTCTTCTAGCCTACCGCTCTCTTTTTTATT	1920
QY	1921	ACTTTTGAAGAAATTCAAATCAAAACAGATCAAAATTAACACGGTGGAGACTGTGCAT	1980
DB	1921	ACTTTTGAAGAAATTCAAATCAAAACAGATCAAAATTAACACGGTGGAGACTGTGCAT	1980
QY	1981	GCTAGTCTCTGAAAGCAATTAATTCGCGCATCCACAGACGTGCTCAGCTTTCATCACCAC	2040
DB	1981	GCTAGTCTCTGAAAGCAATTAATTCGCGCATCCACAGACGTGCTCAGCTTTCATCACCAC	2040
QY	2041	TTTTTTCCTACATACCATGTCCGATGCTTGTGTGATGACACACACACACAGCTTGCTT	2100
DB	2041	TTTTTTCCTACATACCATGTCCGATGCTTGTGTGATGACACACACACACAGCTTGCTT	2100
QY	2101	TGGTTGTGCTTAACAGAGAGAGAGAGACAGACCGATAGCTCTCTCAATTCACATGG	2160
DB	2101	TGGTTGTGCTTAACAGAGAGAGAGAGAGACAGACCGATAGCTCTCTCAATTCACATGG	2160

RESULT 2
US-09-892-635A-44
; Sequence 44, Application US/09892635A
; Publication No. US20030226175A1

GENERAL INFORMATION:
; APPLICANT: May, Gregory D.
; APPLICANT: Clendennen, Stephanie K.
; APPLICANT: Mason, Hugh S.
; APPLICANT: Lim, Miguel A. Gomez
; APPLICANT: Airtzen, Charles J.
; TITLE OF INVENTION: DNA Regulatory Elements Associated with Fruit Development
; FILE REFERENCE: 031998-007
; CURRENT APPLICATION NUMBER: US/09/892,635A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 09/160,351
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: US 60/060,062
; PRIOR FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 2156
; TYPE: DNA
; ORGANISM: Musa acuminata
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 879..956
; OTHER INFORMATION: n = A,T,C or G
US-09-892-635A-44

Query Match 99.0%; Score 2139; DB 13; Length 2156;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2155; Conservative 0; Mismatches 1; Indels 4; Gaps 1;
QY 1 GGATCCCACTTTTAGGAATGGATCTTAAAAATTTAGTTATAGTTTCAAAAGTTAGAAAA 60
Db 1 GGATCCCACTTTTAGGAATGGATCTTAAAAATTTAGTTATAGTTTCAAAAGTTAGAAAA 60
QY 61 TCTTTACCAAGAGCTTTGAGTCACATGATGACATCCCGTGAACCGGTACATGTCCTCGA 120
Db 61 TCTTTACCAAGAGCTTTGAGTCACATGATGACATCCCGTGAACCGGTACATGTCCTCGA 120
QY 121 TGGACTCAGTGTGTTTCATTCGGAAGAGTTTCGAAAGAGTCATGAAAGATTTGATTTTGG 180
Db 121 TGGACTCAGTGTGTTTCATTCGGAAGAGTTTCGAAAGAGTCATGAAAGATTTGATTTTGG 180
QY 181 ATTCTTTCTACTCGGTTGGTGCCTTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAAG 240
Db 181 ATTCTTTCTACTCGGTTGGTGCCTTCATGAGTGACCTCAAGAGTCCCTCCAAATATCAAAAG 240
QY 241 CCGAATCAGAAATGAAATGTGATTTGAATTCATTTTGTCTAATGCACAAAAACAGGSCAT 300
Db 241 CCGAATCAGAAATGAAATGTGATTTGAATTCATTTTGTCTAATGCACAAAAACAGGSCAT 300
QY 301 TCATAGCCTTTGTGTTTAAAGCAAAACATTTCTCTCCGATTCATCCCATTCGCTCATCG 360
Db 301 TCATAGCCTTTGTGTTTAAAGCAAAACATTTCTCTCCGATTCATCCCATTCGCTCATCG 360
QY 361 GAAGAGAAAATTTTGAAATCCATTTTCGAAATAGAACCAAGCTCGAAATCCATCGATG 420
Db 361 GAAGAGAAAATTTTGAAATCCATTTTCGAAATAGAACCAAGCTCGAAATCCATCGATG 420
QY 421 GAAATCAGAGAGATCTCTATGAGTTTCCAAATACATGTAATTCGATCAATTAACATA 480
Db 417 GAAATCAGAGAGATCTCTATGAGTTTCCAAATACATGTAATTCGATCAATTAACATA 476
QY 481 GGTGGATGTGTAATGAAATGACCTCATGSCATCTCTCTGGGTATTAACCAATAT 540
Db 477 GGTGGATGTGTAATGAAATGACCTCATGSCATCTCTCTGGGTATTAACCAATAT 536
QY 541 GAGAGTGAGCTTTGCTCTGATACCAATTTGTAGGATCAGAGTGGCACTAAGAGAGGGGG 600
Db 537 GAGAGTGAGCTTTGCTCTGATACCAATTTGTAGGATCAGAGTGGCACTAAGAGAGGGGG 596
QY 601 GAGTGAATTAGTCAGTGGATTAATAAGTTTAAAAATGAAATCGTAAATACGAG 660
Db 597 GAGTGAATTAGTCAGTGGATTAATAAGTTTAAAAATGAAATCGTAAATACGAG 656

QY 661 AAGATTTTCGTTTAAATAGTAACTTGTAGATGAAACCAAAAGTTAACTAGTGTAAA 720
Db 657 AAGATTTTCGTTTAAATAGTAACTTGTAGATGAAACCAAAAGTTAACTAGTGTAAA 716
QY 721 TAACAAATTCGGGAAAGTAAAGACTCACAATTCAGAGACATACCAATTTAAAGTGT 780
Db 717 TAACAAATTCGGGAAAGTAAAGACTCACAATTCAGAGACATACCAATTTAAAGTGT 776
QY 781 CGGTCAAAATGACCTTACATCCACTTGTGAAGCCCTTTCTGGAAGAGGTCCCAACTTCAC 840
Db 777 CGGTCAAAATGACCTTACATCCACTTGTGAAGCCCTTTCTGGAAGAGGTCCCAACTTCAC 836
QY 841 TAGCAAAATCACTTTGAAGGGAAGGACAAATACCTCTCTTACNACCTTTTACAATGGTTC 900
Db 837 TAGCAAAATCACTTTGAAGGGAAGGACAAATACCTCTCTTACNACCTTTTACAATGGTTC 896
QY 901 ATACTCTTACAAATTTTCAACGAGAAAGAGGAGTGAACATGCAAGCAATTTGAAACAA 960
Db 897 ATACTCTTACAAATTTTCAACGAGAAAGAGGAGTGAACATGCAAGCAATTTGAAACAA 956
QY 961 GACTTGTCTAAAGACTTTTGTCTAAGGCTTTTCTCAATCTATTGCTTCTCAAAAGTTGTA 1020
Db 957 GACTTGTCTAAAGACTTTTGTCTAAGGCTTTTCTCAATCTATTGCTTCTCAAAAGTTGTA 1016
QY 1021 TTCTCTGCTGAGAAATTTAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGTCCA 1080
Db 1017 TTCTCTGCTGAGAAATTTAGGGGTATTTATAGACCCCAAGAGGATTTAAATTTGGGTCCA 1076
QY 1081 AATTTCGAATGCTTCTGGGTTCCCGAGGTTCCCGGTGCCACCGCTCTCAGTGTTCACA 1140
Db 1077 AATTTCGAATGCTTCTGGGTTCCCGAGGTTCCCGGTGCCACCGCTCTCAGTGTTCACA 1136
QY 1141 CTGGACAGTGTACTAGCGGTSCCACCCTCGGCTGTTCGGGTGTGGCGGTGCCACCGC 1200
Db 1137 CTGGACAGTGTACTAGCGGTSCCACCCTCGGCTGTTCGGGTGTGGCGGTGCCACCGC 1196
QY 1201 CTAGACTTTTCTAGCTCAGTGTGATTCCAAATTTGACCCCAAAACAGTTCGAACTCGG 1260
Db 1197 CTAGACTTTTCTAGCTCAGTGTGATTCCAAATTTGACCCCAAAACAGTTCGAACTCGG 1256
QY 1261 GTCCAAATGACCCGTAACCGGATTTAGGATTAACCTTAATCTTAACCTTAATATATG 1320
Db 1257 GTCCAAATGACCCGTAACCGGATTTAGGATTAACCTTAATCTTAACCTTAATATATG 1316
QY 1321 CAAACTACGCAACTGAAATATATAGTCTTAAGCAAGTTTAAACCGGCAAAACGTCGAGTCT 1380
Db 1317 CAAACTACGCAACTGAAATATATAGTCTTAAGCAAGTTTAAACCGGCAAAACGTCGAGTCT 1376
QY 1381 TCTTCCGGGATCTTTTCGGGAGACTTCTGATATACCTTTGGATTTCTTTCAGCGGACTCC 1440
Db 1377 TCTTCCGGGATCTTTTCGGGAGACTTCTGATATACCTTTGGATTTCTTTCAGCGGACTCC 1436
QY 1441 TAGTAGGTTCCGATCTGTGGCGAGTTTAGCGAGTAGCGGACCTTCTCGGTGATCTCC 1500
Db 1437 TAGTAGGTTCCGATCTGTGGCGAGTTTAGCGAGTAGCGGACCTTCTCGGTGATCTCC 1496
QY 1501 GCAAAACCGCGATGATCTCTTCGGCAGACTTTTCAAAACCTTCGACAAAGTCCCGGATTTCT 1560
Db 1497 GCAAAACCGCGATGATCTCTTCGGCAGACTTTTCAAAACCTTCGACAAAGTCCCGGATTTCT 1556
QY 1561 TCTCGGTTGGTTCCGACAGCTCTTAACGAAATTTTCGAAATCTTTCGAAATGTCATCGAA 1620
Db 1557 TCTCGGTTGGTTCCGACAGCTCTTAACGAAATTTTCGAAATCTTTCGAAATGTCATCGAA 1616
QY 1621 CTTGACTCCGTTAGGCTCTTTTATATTTTCAGGCTATCATAGTTAATCTTACATCTTA 1680
Db 1617 CTTGACTCCGTTAGGCTCTTTTATATTTTCAGGCTATCATAGTTAATCTTACATCTTA 1676
QY 1681 ACTCAATAATATGATTTAGATTAATTAACCCATCAATTTGATTTTCATCATCAAAATTCGAC 1740
Db 1677 ACTCAATAATATGATTTAGATTAATTAACCCATCAATTTGATTTTCATCATCAAAATTCGAC 1736

QY	1261	GTCCAAATTGACCGGTAAACCGGATTATAGGATTAAACCCCTTAATCTCTAAACCCCTAAATTATATG	1320
DB	1257	GTCCAAATTGACCGGTAAACCGGATTATAGGATTAAACCCCTTAATCTCTAAACCCCTAAATTATATG	1316
QY	1321	CAAACTAGCGAACTGAAATAATAGTCTTAAGCAAGTTTTAAACCGGCAAAAGTGCAGTCT	1380
DB	1317	CAAACTAGCGAACTGAAATAATAGTCTTAAGCAAGTTTTAAACCGGCAAAAGTGCAGTCT	1376
QY	1381	TCTTCCGGCGGATCTTTTCGGCAGACTCTGTGATATACCTTTGGATTCTTCTAGCGGACTCC	1440
DB	1377	TCTTCCGGCGATCTTTTCGGCAGACTCTGTGATATACCTTTGGATTCTTCTAGCGGACTCC	1436
QY	1441	TAGTAGGTCCTCGAFTCTGTGGCGAGTTTAGCGAGTAGCCGAACTTCTCGGTGAUCTCC	1500
DB	1437	TAGTAGGTCCTCGAFTCTGTGGCGAGTTTAGCGAGTAGCCGAACTTCTCGGTGAUCTCC	1496
QY	1501	GCAACCGCGCATGATCTTTTCGGCAGACTTTTGGAAACTTTCGAAAGTTCGCAAGTCCCGATTCT	1560
DB	1497	GCAACCGCGCATGATCTTTTCGGCAGACTTTTGGAAACTTTCGCAAGTTCGCGATTCT	1556
QY	1561	TCTCGGTTGGTTCGACAGCATCTCTAACGAAACTTCGGACTCTCTGAAATGTCATCGAA	1620
DB	1557	TCTCGGTTGGTTCGACAGCATCTCTAACGAAACTTCGGACTCTCTGAAATGTCATCGAA	1616
QY	1621	CTTGACTCCGGTAGGCTTGCTTTATATTTTCAGGCTATCATAGTTAATCTCATACTCTTA	1680
DB	1617	CTTGACTCCGGTAGGCTTGCTTTATATTTTCAGGCTATCATAGTTAATCTCATACTCTTA	1676
QY	1681	ACTCAATAATAGATTAGATTAAATTAACCGATCAATGATTTTCATCATCAAAATTCGAC	1740
DB	1677	ACTCAATAATAGATTAGATTAAATTAACCGATCAATGATTTTCATCATCAAAATTCGAC	1736
QY	1741	ATTCAACAAACATCCGTACTCAATAACCCATCAGGCTATAGTTACGTTGACTACTCTGT	1800
DB	1737	ATTCAACAAACATCCGTACTCAATAACCCATCAGGCTATAGTTACGTTGACTACTCTGT	1796
QY	1801	GATCCGTACGTGAAGTTAGCGAGTCATGATCCAGGTCGTCTCATTTATGGCCGAAACG	1860
DB	1797	GATCCGTACGTGAAGTTAGCGAGTCATGATCCAGGTCGTCTCATTTATGGCCGAAACG	1856
QY	1861	TATCCCTTATCCAAATCCAGTCTCTCAACTCTTCAGGCTACCGGTCCTTTTTTTTAT	1920
DB	1857	TATCCCTTATCCAAATCCAGTCTCTCAACTCTTCAGGCTACCGGTCCTTTTTTTTAT	1916
QY	1921	ACTTTTGAAAGAAATTCAAATCAAAACAGATACAAATAATACACGGTGAGACACTGTGCAT	1980
DB	1917	ACTTTTGAAAGAAATTCAAATCAAAACAGATACAAATAATACACGGTGAGACACTGTGCAT	1976
QY	1981	GCTAGTCTCTGGAAAGCATTAATTCGGCGATCCACAGACGTCTGAGTTTCATCACCAC	2040
DB	1977	GCTAGTCTCTGGAAAGCATTAATTCGGCGATCCACAGACGTCTGAGTTTCATCACCAC	2036
QY	2041	TTTTTCTCATCATACCATGTCGCTTTGTTGATGACAGACCAACGCTTGCT	2099
DB	2037	TTTTTCTCATCATACCATGTCGCTTTGTTGATGACAGACCAACGCTTGCT	2096
QY	2100	TTGGTTGTGCTTAACAGAGAGAGAGAGACAGACCGATAGCTTCCTCATTCACCATG	2159
DB	2097	TTGGTTGTGCTTAACAGAGAGAGAGAGACAGACCGATAGCTTCCTCATTCACCATG	2155
QY	2160	G 2160	2162
DB	2153	G 2153	2157

```

RESULT 4
US-10-298-122-3
; Sequence 3, Application US/10298122
; Publication No. US2003022114A1
; GENERAL INFORMATION:
; APPLICANT: Gmitter, Frederick G
; APPLICANT: Deng, Zhanao

```

```

: APPLICANT: Zhang, Hongbin
: TITLE OF INVENTION: CITRUS TRISTEZA VIRUS AND METHODS OF USE
: FILE REFERENCE: 5853-220
: CURRENT APPLICATION NUMBER: US/10/298,122
: CURRENT FILING DATE: 2002-11-15
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3
: LENGTH: 7833
: TYPE: DNA
: ORGANISM: artificial
: FEATURE:
: OTHER INFORMATION: Poncirus trifoliata and citrus grandis hybrid
US-10-298-122-3

```

Query Match	5.7%;	Score 122.2;	DB 16;	Length 78333;
Best Local Similarity	59.8%;	Pred. No. 7e-21;		
Matches 222;	Conservative 0;	Mismatches 148;	Indels 1;	Gaps 1;

QY	6	CGAACCTTTTAGGAAGTGCATCTTAAAAATTTTAGTTTAAAGTTCAAAAGTTAGAAAAATCTTTT	65
Db	47226	CCACTCCTTAGGTAGAGATCTAATGATTTTCTGTACTTTTTCGGTATTGTGAAAAGTCTT	47285
QY	66	ACCAAGAGCTTTTGAGTCCATTGATGACATCCCGTGAAC-GGTGTACATGTTCTCCGATGGA	124
Db	47286	TCTAGGGCTCTAGAGTGTTTCACTATGTCCTGTAATCTAGTATACATAGAGTACACATT	47345
QY	125	CTCAGTTGGTTTCATTTCGAAAAGTTCGAAAGAGTGCATTAAGATATTGATTTTGGATTC	184
Db	47346	TTCAATTTTGTCCATTGGAACAAATTCGTATTGTGCAGTGTATCTACTTATTTTAGATT	47405
QY	185	TTTCACTCGGTTGGTGGCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAGCGA	244
Db	47406	TTTCACTTGGTTTCGTACCTTCATAGACAACTCAAGTTTGTGCCAAATCTCATTAGCACT	47465
QY	245	ATCACAAATTGAAATGTGATTGAATTCATTTTGTCTAATGCACAAAACAGGGCATTCAT	304
Db	47466	TTTCGCAACTAGACACTCTATGAAACTCTTACTTATCTAGTGCACAGAACAAAGGCATTCAT	47525
QY	305	AGCCVTTTGTGTTTAAAGCAAAAAACATCTTCTCCGATTTCATCCCATTCGCTCATCGGAAG	364
Db	47526	GGCCVTTGGAATTTAGAGAAGCTTTTCTCTTATGAAATTCATTCCAATCCCGTGAAGGTTT	47585
QY	365	AGAAAAATTTTT	375
Db	47586	TGGAATATCTT	47596

RESULT 5

```

US-10-424-599-108338
; Sequence 108338, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 108338
; LENGTH: 2760
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT3847_68844C.1
US-10-424-599-108338

```

Query Match 3.8%; Score 81.4; DB 13; Length 2760;
Best Local Similarity 53.8%; Pred. No. 9.7e-11;

Matches 189; Conservative 0; Mismatches 161; Indels 1; Gaps 1;
QY 1 GGATCCCAACTTTTAGAATGGATCTTAAATTTTGTATTAAGTTCAAGTTAGAAAA 60
Db 1829 GGTGCGCATCTCTACTTAAACATCTTAAACACCTTGTATTATAGATCTCTATTTTGAAT 1888
QY 61 TCTTTACCAAGAGCTTTGAGTCCATTCATGACATCCCTGAAACGGTGTACATGTCCTCG- 119
Db 1889 TCTTTGCTAAGCTGTAGATGATTACTATATGTGTAATCTCTTTTGCATCTCTGA 1948
QY 120 ATGACTCAGTCTGTTTCATTGGGAAAAGTTTCCGAAGAGTGTCATAAGAAATATTGATTG 179
Db 1949 ATATTTTCATTTGCAATTCATCTTAAATTAATTCATCTCAAGTTAGTTCATTTATCCTA 2008
QY 180 GATCTTTTCACTCGGTTGGTCCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAA 239
Db 2009 GATCTTTTAAACATTTGATGTTCTTCTCATGTGTTAATCGAAGAGTGTCCTCCATTTCTCTTA 2068
QY 240 GCCGAATCACAATTAAGAAATGATGATTGAATTCATTTTGTCTAATGACAAAAACAGGGCA 299
Db 2069 GCATCTTACAATTTGAACCCCTCGAAATATTCATCCATTCCTAGGAGATGTTATTATG 2128
QY 300 TTCTAGCTCTTGTGTTTAAAGCAAAAACATCTTCTCCGATTCATCCCAT 350
Db 2129 TTTTGGCTTTAAGTTGATTGTACTCTTTTCTATCTCTTCAGACCAT 2179

RESULT 6

US-10-298-122-2
; Sequence 2, Application US/10298122
; Publication No. US2003022141A1
; GENERAL INFORMATION:
; APPLICANT: Gmitter, Frederick G
; APPLICANT: Zhang, Zhanao
; APPLICANT: Zhang, Hongbin
; TITLE OF INVENTION: CITRUS TRISTEZA VIRUS AND METHODS OF USE
; FILE REFERENCE: 5853-220
; CURRENT APPLICATION NUMBER: US/10/298,122
; CURRENT FILING DATE: 2002-11-15
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 58965
; TYPE: DNA
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Poncirus trifoliata and Citrus grandis hybrid
; NAME/KEY: misc feature
; LOCATION: (2145)..(2320)
; OTHER INFORMATION: n denotes unsequenced nucleotides
US-10-298-122-2

Query Match 3.8%; Score 81.2; DB 16; Length 58965;
Best Local Similarity 57.7%; Pred. No. 8.1e-10;
Matches 188; Conservative 0; Mismatches 128; Indels 10; Gaps 2;
QY 181 ATTCTTCACTCGGTTGGTCCCTTCATGAGTCACTCAAGAGTCTCCAAATATCAAAAG 240
Db 38730 ATTCTCTCTCTTGTAGTCTCTCATACACAAATTTTGAGTTTCTCCAAATTTTCATAAG 38789
QY 241 CCGATATCAATTCAGAAATGATGATTGAATTCATTTTGTCTAATGCACAAAACAGGCAT 300
Db 38790 CATTAGACAAATTCAGAACTCTATGAATTCCTTTTATCTAGGACACAAATAAGAT 38849
QY 301 TCATAGCCTTTGTGTTTAAAGCAAAAACATCTTCTCCGATTCATCCCATTCGCTCATCG 360
Db 38850 ICATAGCTTTGAAATTTAAAGACATATTTTCTATCTAACTCACTTAAAGCATCTTTC 38909
QY 361 GAAGGAAATTTTGAATTCATTTTCGACATAGACCAAGCTCGAATCCATGCATG 420
Db 38910 TAGGA-----ATGAGTAAATCATCACAAACAACTTCCCAATTTTCATAATCTAAAGCTT 38963

QY 421 GAAATGAGGAAGATCCTCATATGAGTTTCCAAATACATGTAATTCGACTCATTAACAATA 480
Db 38964 GTAA-----ATAAATTCATCTTAGTTTTCCTAATAAGGATAATCATTTCATCTAAAAAT 39019
QY 481 GGTGGATGTGTAATGAATGACCCCTC 506
Db 39020 GGAAGTCTTGTGATAGATTGCTCTTC 39045
RESULT 7
US-09-966-881-48
; Sequence 48, Application US/09966881
; Patent No. US20020120960A1
; GENERAL INFORMATION:
; APPLICANT: Seymour, Graham
; Bird, Colin
; Medina-Suarez, Rosybel
; TITLE OF INVENTION: Genetic control Of Fruit Ripening
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zeneca Ag Products Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/966,881
; FILING DATE: 28-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,860
; FILING DATE: 29-Mar-1999
; APPLICATION NUMBER: GB 9618862.8
; FILING DATE: 10-SEP-1996
; APPLICATION NUMBER: GB 9708366.1
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: PCT/GB97/02424
; FILING DATE: 08-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hohenschutz, Liza D.
; REGISTRATION NUMBER: 33,712
; REFERENCE/DOCKET NUMBER: SEE 50183/UST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 886-1699
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 809 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; CLONE: U-D86
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-09-966-881-48
Query Match 3.4%; Score 73.4; DB 9; Length 809;
Best Local Similarity 98.7%; Pred. No. 6.5e-09;
Matches 74; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2086 CCACAAGCTTGCTTTGTTGTCCTTAACAGAGAGAGAGAGACAGACCGATAGCCT 2145
Db 9 CCACAAGCTTGCTTTGTTGTCCTTAACAGAGAGAGAGAGACAGACCGATAGCCT 68
QY 2146 CCTCATTCCACCATGG 2160
Db 69 CCTCATTCCATATGG 83

RESULT 8
US-10-334-703-3/c
; Sequence 3, Application US/10334703
; Publication No. US2003015451A1
; GENERAL INFORMATION:
; APPLICANT: Laten, Howard M.
; TITLE OF INVENTION: PLANT RETROVIRAL POLYNUCLEOTIDES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, Gerstein and Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/334,703
; FILING DATE: 20-Dec-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Wrona, Thomas J.
; REGISTRATION NUMBER: 44,410
; REFERENCE/DOCKET NUMBER: 27013/33214C US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2417 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-334-703-3

Query Match 2.9%; Score 62.8; DB 15; Length 2417;
Best Local Similarity 50.6%; Pred. No. 1e-05;
Matches 177; Conservative 0; Mismatches 172; Indels 1; Gaps 1;
QY 4 TCCCACTTTTAGGAATGGATCTTAAAAATTTTAGTTTATAAGTTCAAAGTTAGAAAAATCT 63
Db TCAAATCTCTTAGGCAAGGATCTGAGGATCTTTCTCACCAGCTTTTCATCAGTCATTCCT 727
QY 64 TTACCAAGAGCTTTGAGTCAATGATGACATC-CGTGAACGGGTGATCATGTCCTCCGATG 122
Db TCTCCCAAGCAGTGAAGCAATTTGCAAAATTTCAAGAAATGTTCAATGTCGAAAGTCATGAATA 667
QY 123 GACTCACTTTGTTTCATTCGGAAGAGTTCGAAAGAGTGCATAGATAATTTGATTTGGAT 182
Db CACTCTTCCTCTTCATCTCAGATTTTCAATTTTGTGGCCAAATAGTTGCAATCTGGAC 607
QY 183 TCTTTCACCTGGTGTGGCTTCATGATGACATCAAGAGTCCTCCAAATATCAAAAGCC 242
Db ATCTTCACCTTGGAGGTCTCTTCATGATGTTTTCAGGATCTCCCATGATCCTTGGCC 547
QY 243 GAATCAAAATGAAATGATTAATTTCTTAATGCAAAACAGGGCATTC 302
Db ACTGTGCATGTGTGATGATCTGAAGATATTTCTGTCAACTCCATTCATTAATAGAGCATTC 487
QY 303 ATAGCCTTTGTGTTTAAAGCAAAACATTTCTTCCGATTCATCCCATTC 352
Db AAAGCTTTGGAGTTTCCAAGTCCCAATTCGTCTTCTTTAGTCCAGTC 437

RESULT 9
US-10-396-122-3/c
; Sequence 3, Application US/10396122
; Publication No. US20030221222A1
; GENERAL INFORMATION:
; APPLICANT: Laten, Howard
; TITLE OF INVENTION: PLANT RETROVIRAL POLYNUCLEOTIDES AND METHODS FOR USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/396,122
; CURRENT FILING DATE: 2003-03-25
; PRIOR APPLICATION NUMBER: US 60/367,302
; PRIOR FILING DATE: 2002-03-25
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 2417
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: SIRE 1 cDNA clone
US-10-396-122-3

Query Match 2.9%; Score 62.8; DB 16; Length 2417;
Best Local Similarity 50.6%; Pred. No. 1e-05;
Matches 177; Conservative 0; Mismatches 172; Indels 1; Gaps 1;
QY 4 TCCCACTTTTAGGAATGGATCTTAAAAATTTTAGTTTATAAGTTCAAAGTTAGAAAAATCT 63
Db TCAAATCTCTTAGGCAAGGATCTGAGGATCTTTCTCACCAGCTTTTCATCAGTCATTCCT 727
QY 64 TTACCAAGAGCTTTGAGTCCATGATGACATC-CGTGAACGGGTGATCATGTTTGGAT 122
Db TCTCCCAAGCAGTGAAGCAATTTGCAAAATTTCAAGAAATGTTCAATGTCGAAAGTCATGAATA 667
QY 123 GACTCACTTTGTTTCATTCGGAAGAGTTCGAAAGAGTGCATAGATAATTTGATTTGGAT 182
Db CACTCTTCCTCTTCATCTCAGATTTTCAATTTTGTGGCCAAATAGTTGCAATCTGGAC 607
QY 183 TCTTTCACCTGGTGTGGCTTCATGATGACATCAAGAGTCCTCCAAATATCAAAAGCC 242
Db ATCTTCACCTTGGAGGTCTCTTCATGATGTTTTCAGGATCTCCCATGATCCTTGGCC 547
QY 243 GAATCAAAATGAAATGATTAATTTCTTAATGCAAAACAGGGCATTC 302
Db ACTGTGCATGTGTGATGATCTGAAGATATTTCTGTCAACTCCATTCATTAATAGAGCATTC 487
QY 303 ATAGCCTTTGTGTTTAAAGCAAAACATTTCTTCCGATTCATCCCATTC 352
Db AAAGCTTTGGAGTTTCCAAGTCCCAATTCGTCTTCTTTAGTCCAGTC 437

RESULT 10
US-10-424-599-108650/c
; Sequence 108650, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 108650
; LENGTH: 2801
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_69124C.1

US-10-424-599-108650	
Query Match 2.9%; Score 62.8; DB 13; Length 2801;	
Best Local Similarity 50.6%; Pred. No. 1.1e-05;	
Matches 177; Conservative 0; Mismatches 172; Indels 1; Gaps 1;	
QY	4 TCCCAACTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAGAAAGTTAGAAAATCT 63
Db	1093 TCAAAATCTTTAGCAAGGATCTGAGGATCTTTCTCCAGCTTTTCATCAGTCATCTT 1034
QY	64 TTACCAAGAGCTTTGAGTCCATTCGACATC-CGTGAAACGGTGTACATGTCTCCGATG 122
Db	1033 TCTCCCAAGCAGTGCAGCATTTGCCAATTCAGAAATGTTCAATGTGAAAGTCATGAATA 974
QY	123 GACTCACTTGGTTTCATTCGAAAGTTCGAAAGAGTGATCAAGAAATTTGATTTGGAT 182
Db	973 CACTCTTCCCTCCTTCATCTTCAGATTTTCGAAATTTTGGGCCAATAGTTGCAATCTGGAC 914
QY	183 TCTTTCACTCGTGTGGTCTTCATGAGTGACATC-CGTGAAACGGTGTACATGTCTCCGATG 122
Db	913 ATCTTCACCTTTGGAGGTTCTTCATGAGTGATTTTTCAGGATCTCCCATGCAATCTTGCC 854
QY	243 GAATCACAAAATGAAATGTAATGTAATCAATTTTGTCTAATGCACAAAACAGGGCAATC 302
Db	853 ACTGTGATGTGTTGATCAGTCTGAGATATTTCTGTCACTTCCATTGAATGATGACATTC 794
QY	303 ATAGCTTTGTGTTTAAAGCAAAAACATCTTCTCCGATTCATCCCATC 352
Db	793 AAAGCTTTGGAGTTTCCAAAGTGCAATTCGTCTTCTTCTTTAGTCCAGTC 744
RESULT 11	
US-10-396-122-93/c	
; Sequence 93, Application US/10396122	
; Publication No. US2003021222A1	
; GENERAL INFORMATION:	
; APPLICANT: Laten, Howard	
; TITLE OF INVENTION: PLANT RETROVIRAL POLYNUCLEOTIDES AND METHODS FOR USE THEREOF	
; FILE REFERENCE: 27013/39234	
; CURRENT APPLICATION NUMBER: US/10/396,122	
; CURRENT FILING DATE: 2003-03-25	
; PRIOR APPLICATION NUMBER: US 60/367,302	
; PRIOR FILING DATE: 2002-03-25	
; NUMBER OF SEQ ID NOS: 105	
; SOFTWARE: Patent in version 3.0	
; SEQ ID NO 93	
; LENGTH: 9399	
; TYPE: DNA	
; ORGANISM: Glycine max	
; FEATURE:	
; NAME/KEY: misc_feature	
; OTHER INFORMATION: Soybean retroelement SIRE1 9	
; FEATURE:	
; NAME/KEY: CDS	
; LOCATION: (1215)..(5942)	
; OTHER INFORMATION: gag protease reverse transcriptase RNAase H polyprotein	
; FEATURE:	
; NAME/KEY: CDS	
; LOCATION: (5946)..(7985)	
; OTHER INFORMATION: envelope protein	
; FEATURE:	
; NAME/KEY: LTR	
; LOCATION: (29)..(1156)	
; FEATURE:	
; NAME/KEY: LTR	
; LOCATION: (8606)..(9361)	
US-10-396-122-93	
Query Match 2.9%; Score 62.8; DB 16; Length 9399;	
Best Local Similarity 50.6%; Pred. No. 2.5e-05;	
Matches 177; Conservative 0; Mismatches 172; Indels 1; Gaps 1;	
QY	4 TCCCAACTTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAGAAAGTTAGAAAATCT 63

US-10-437-963-9498	
; Sequence 9498, Application US/10437963	
; Publication No. US2004012343A1	
; GENERAL INFORMATION:	
; APPLICANT: La Rosa, Thomas J.	
; APPLICANT: Kovalic, David K.	
; APPLICANT: Zhou, Yihua	
; APPLICANT: Cao, Yongwei	
; APPLICANT: Wu, Wei	
; APPLICANT: Boukharov, Andrey A.	
; APPLICANT: Barbazuk, Brad	
; APPLICANT: Li, Ping	
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With	
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement	
; FILE REFERENCE: 38-21(53221)B	
; CURRENT APPLICATION NUMBER: US/10/437,963	
; CURRENT FILING DATE: 2003-05-14	
; NUMBER OF SEQ ID NOS: 204966	
; SEQ ID NO 9498	
; LENGTH: 5371	
; TYPE: DNA	
; ORGANISM: Oryza sativa	
; FEATURE:	
; OTHER INFORMATION: Clone ID: PAT_MRT4530_1590C.1	
US-10-437-963-9498	
Query Match 2.8%; Score 61.2; DB 17; Length 5371;	
Best Local Similarity 52.3%; Pred. No. 4.7e-05;	
Matches 158; Conservative 0; Mismatches 145; Indels 1; Gaps 1;	
QY	4 TCCCAACTTTTAGGAATGGATCTTAAATTTTAGTTATAAGTTCAGAAAGTTAGAAAATCT 63
Db	4864 TCATAATTTCTCGGTAGAGCCCTAAGCATCTTTTGGCAACCTCAAAATCGGTAGATT 4923
QY	64 TTACCAAGAGCTTTGAGTCCATTCGACATCCGCTGAAAC-GGTGTACATGTCTCCGATG 122
Db	4924 GCCCCAAGCCCTTGAGATCATTCACAATAACATTGAGCCCTCCCATATATGTCATTCAG 4983
QY	123 GACTCACTTGGTTTCATTCGAAAGTTCGAAAGAGTGATCAAGAAATTTGATTTGGAT 182
Db	4984 CTCTCATGAGCAACATGGAGATCTCTCATATATAGATCTTCAGAAATGAAGCTTGCA 5043
QY	183 TCTTTCACTCGGTGGTGCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAGCC 242
Db	5044 TCTTTGTACTCGCTTGTGCCCTCATGGATCTCCGCCAACTTGTTCCAAATCTCATATGCG 5103
QY	243 GAATCACAAAATGAAATGTAATGTAATTCATTTTGTCTAATGCACAAAACAGGGCATTC 302

Db 5104 GTCTCAAGGTTGCTCACTCTATCAAACTCCCTCTGGCTCAAGAGATTGAACAGGCATTC 5163

Qy 303 AT 304

Db 5164 AT 5165

RESULT 13

US-10-396-122-87/c

; Sequence 87, Application US/10396122

; Publication No. US20030221222A1

; GENERAL INFORMATION:

; APPLICANT: Laten, Howard

; TITLE OF INVENTION: PLANT RETROVIRAL POLYNUCLEOTIDES AND METHODS FOR USE THEREOF

; FILE REFERENCE: 27013/39234

; CURRENT APPLICATION NUMBER: US/10/396,122

; CURRENT FILING DATE: 2003-03-25

; PRIOR APPLICATION NUMBER: US 60/367,302

; PRIOR FILING DATE: 2002-03-25

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 87

; LENGTH: 9072

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: SIRE1 7

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1264)..(5991)

; OTHER INFORMATION: gag protease integrase reverse transcriptase RNAase H polyprotein

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (5995)..(8040)

; OTHER INFORMATION: envelope protein

; FEATURE:

; NAME/KEY: LTR

; LOCATION: (58)..(1205)

; FEATURE:

; NAME/KEY: LTR

; LOCATION: (8677)..(9072)

US-10-396-122-87

Query Match 2.8%; Score 59.6; DB 16; Length 9072;

Best Local Similarity 50.0%; Pred. No. 0.00018;

Matches 175; Conservative 0; Mismatches 174; Indels 1; Gaps 1;

Qy 4 TCCCACTTTTAGGAATGGATCTTAAATTTTAGTTTATAGTTTCAAAAGTTAGAAAAATCT 63

Db 1814 TCAAAATCTCTTAGGCAAGGATCTGAGGATCTTTCTCACCAGCTTTTCATCTGTTATCCCTC 1755

Qy 64 TTACCAAGAGCTTTGAGTCCATTTGATGACATC-CGTGAAACGGTGACATGCTCCGATG 122

Db 1754 TCTCCCAAGCAGTGAAGCAATTTGCAATTTTCAAGAAATGTTCAATGTTGGAAGTCAATGAATA 1695

Qy 123 GACTCACTTGGTTTCAATCGAAAAAGTTTCAAGAGTGCATAGAATATTTGATTTGGAT 182

Db 1694 CACTCTCTCTCTTCACTTTCAGATTTTCAATTTTGTAGCCAAAGTTGCAATCTGGAC 1635

Qy 183 TCTTTCACTCGGTGTGGCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAGCC 242

Db 1634 ATCTTCACTTTGGAGGTTCTTCATGAGTGCTTTTCAGGATCTCGCATGCACTTTGGCC 1575

Qy 243 GAATCAAAATGAATGATGAATTTGTTGTTCTTAATGCAAAACAGGGCAATTC 302

Db 1574 ACTGTGCAAGTGTGTGATGATGATGTTCTTGTCAACTCAATTGAATAGTGCATTC 1515

Qy 303 ATAGCCCTTTGTTTAAAGCAAAACATTTCTCTCCGATTCATCCCATTC 352

Db 1514 AAGCTTTGGAGTTTCCAAAGTGCCAAATCGTCTCTCTTTTGTAGTCCAGTC 1465

RESULT 14

US-10-437-963-10782/c

; Sequence 10782, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 10782

; LENGTH: 5016

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_1706C.1

US-10-437-963-10782

Query Match 2.8%; Score 59.4; DB 17; Length 5016;

Best Local Similarity 48.9%; Pred. No. 0.00014;

Matches 251; Conservative 1; Mismatches 247; Indels 14; Gaps 3;

Qy 4 TCCCACTTTTAGGAATGGATCTTAAATTTTAGTTTATAGTTTCAAAAGTTAGAAAAATCT 63

Db 500 TCATACTTCTCCGGTAGGGCTCTAAGCATCTTTTGGGCAACCTCAAGATCGGTCTAGTTT 441

Qy 64 TTACCAAGAGCTTTGAGTCCATTTGATGACATCGTGAAC-CGTGTACATGTTCTCCGATG 122

Db 440 GCCCAAGCCCTTGAGATCATTCATGACATGACATGAGCCTCCCATACATGTCATTCACA 381

Qy 123 GACTCACTTGGTTTCAATCGAAAAAGTTTCAAGAGTGCATAAGAATATTTGATTTGGAT 182

Db 380 CTCTCATGAGGCAACATGGAGAATGCTCATATTTGGATCTTGAGGAATGAAGCTTGGCA 321

Qy 183 TCTTTCACTCGGTGTGGCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAGCC 242

Db 320 TCTTGTACTCACTTTGATCCCTCATGATCTCGGCAACTTGTTCCAAATCTCATATGCG 261

Qy 243 GAATCAAAATTCGAAATGATTTGAATTCATTTTCTTAATGCAAAACAGGGCAATTC 302

Db 260 GTCTCAAGGTTGCTCACTCTATCGAACTCTCTTTGGCTCAAGAGTTGAACAGGCATTC 201

Qy 303 ATAGCCCTTTGTTTAAAGCAAAACATTTCTTCTCGGATTCATCCCATTCGCTCATCGGA 362

Db 200 ATGGC-----TTGGCAATTGAGTTGGAGTTACGGTGATCACTCTCCGTCAGGGT 150

Qy 363 AGAGAAAAATTTTGAATCCATTTTCGCAATAGACCAAGCTCGAAATTCATGCAATGA 422

Db 149 GTGCCGGTGATATCAAGGCTTACATCCCAATACTCCAAATATGGAA-----GCTCATAGC 94

Qy 423 AATGAGGAAGATCTTCATATGATTTTCCAAATACATGTAATTCGACTCATTAACACATAG 482

Db 93 TTTGAGGTGAGTAGACATTTTAATTTTCCAGAGTGGAGTAGTTTGTGCCATTAAACATGGA 34

Qy 483 TGGATGTGTAATGAATGACCCCTCATGCSCTAT 515

Db 33 AGCCTCCCTACATGTTCCACCTCGTTCGACAT 1

RESULT 15

US-10-437-963-44132/c

; Sequence 44132, Application US/10437963

; Publication No. US20040123343A1

; GENERAL INFORMATION:

us-09-892-635a-45.rnpb

Search completed: September 18, 2004, 18:55:10
Job time : 971.897 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2004, 17:55:10 ; Search time 4973.1 Seconds
(without alignments)
12970.230 Million cell updates/sec

Title: US-09-892-635A-45
Perfect score: 2160
Sequence: 1 ggatcccaacttttagaat.....agctctcattcaccatgg 2160

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_estl:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_hiv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gsl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110.4	5.1	500	14	CD474210 nad03-10m
2	96.6	4.5	790	28	BH021865 GH_MBB000
3	92.6	4.3	977	29	CG823740 SOYDB38TH
4	92	4.3	832	29	CG818965 SOYED06TV

C 5	91.4	4.2	821	29	CG817604
C 6	89.8	4.2	776	29	CG822306
C 7	83	3.8	732	28	BH794864
C 8	83	3.8	791	29	CG824360
C 9	81.8	3.8	890	29	CG817563
C 10	81.8	3.8	927	29	CG820967
C 11	81.4	3.8	733	29	CG815315
C 12	81.4	3.8	872	29	CG817875
C 13	81.2	3.8	295	14	CB340076
C 14	81.2	3.8	795	29	CG816026
C 15	81.2	3.8	914	29	CG822083
C 16	80.6	3.7	964	28	CC412052
C 17	79.8	3.7	396	29	AG336150
C 18	79.8	3.7	746	29	CG821955
C 19	79.8	3.7	760	29	CG817507
C 20	79.8	3.7	862	29	CG813886
C 21	79.4	3.7	699	12	BI698341
C 22	79	3.7	750	29	CG816271
C 23	78.2	3.6	607	14	CF920770
C 24	78.2	3.6	709	29	CG820051
C 25	78.2	3.6	846	29	CG819170
C 26	78.2	3.6	941	29	CG818537
C 27	78.2	3.6	986	29	CG886466
C 28	77.4	3.6	787	29	CC804388
C 29	77.4	3.6	845	28	BZ648290
C 30	76.8	3.6	352	29	CG816199
C 31	76.8	3.6	722	12	BI678830
C 32	76.8	3.6	803	28	BZ966238
C 33	76.6	3.5	608	28	BZ778766
C 34	76.6	3.5	860	29	CG823490
C 35	76.4	3.5	767	29	CG795738
C 36	76.4	3.5	792	29	CG900423
C 37	76.4	3.5	970	29	CG878233
C 38	76.2	3.5	873	29	CG227901
C 39	75.8	3.5	707	28	BZ527252
C 40	75.8	3.5	856	29	CG734878
C 41	75.4	3.5	383	28	BH780259
C 42	75	3.5	283	14	CB340514
C 43	75	3.5	770	29	CG824885
C 44	75	3.5	787	29	CG678004
C 45	75	3.5	826	29	CG677989

ALIGNMENTS

RESULT 1	CD474210	500 bp	mrna	linear	EST 04-JUN-2003
LOCUS	CD474210	nad03-10ms1-g09	Nuphar advena	cdna	clone nad03-10ms1-g09 5',
DEFINITION	CD474210	mrna sequence.			
ACCESSION	CD474210				
VERSION	CD474210.1	GI:31395478			
KEYWORDS	EST.				
SOURCE	Nuphar advena				
ORGANISM	Nuphar advena				
REFERENCE	1 (bases 1 to 500)				
AUTHORS	dePamphilis, C., Carlson, J., Ma, H., Frohlich, M., Tanksley, S., Leebens-Mack, J., Field, D., Arrington, J., Zahn, L., Kong, H., Druckemiller, M., Landherr, L., Hu, Y., Ilut, D., Wall, K., Plock, S., Chioresan, S., Albert, V., Doyle, J., Miller, W., Oppenheimer, D., Soltis, D., Soltis, P. and Theissen, G.				
TITLE	Generation of ESTs from early flower buds of Nuphar advena				
JOURNAL	Unpublished (2002)				
COMMENT	Contact: Claude dePamphilis or James Leebens-Mack Mueller Laboratory Penn State University 208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn State University, University Park, PA 16802, USA Tel: 814 863 6413				

H07C21:WTP9G4, genomic survey sequence.

C823740 GI:38284664

GSS.

Glycine max (soybean)

Glycine max

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eusoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 977)

Shultz,J., Mekeem,K., Shetty,J., Town,C.D., Koo,H., Potter,J., Wakefield,K., Zhang,H., Wu,C. and Lightfoot,D.A.

End sequencing of BACs comprising a provisional minimal tiling path from a fingerprint physical map of soybean (Glycine max) cultivar Forrest

Unpublished (2003)

Other GSSs: SOYDB38TV

Contact: Chris Town, J. L. Shultz and D. A. Lightfoot

The Center of Excellence in Soybean Research, Teaching and Outreach, Southern Illinois University at Carbondale and Plant Genomics, The Institute for Genomic Research

Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415, USA and 9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 618 453 1797

Email: ga4082@siu.edu, jlschultz@siu.edu, cdtown@stigr.org (URL: http://bioinformatics.siu.edu)

Clones approximating a minimum tiling path were re-arrayed from the library master plates prior to sequencing.

For purposes of clone identification each clone name is a concatenation of the original clone location and its new location in the re-arrayed sequencing plates.

Seq primer: AATTAACCTCTACTAAGGG

Class: BAC ends.

Location/Qualifiers

1..977

/organism="Glycine max"

/mol_type="genomic DNA"

/cultivar="Forrest"

/db_xref="taxon:3847"

/clone="H07C21:WTP9G4"

/clone_lib="LargeInsertSoybeanGenLib"

/note="Organ: Leaves: Vector: pCLD04541 (pBELOBAC11 EcoRI clones): Site 1: BatX1; Soybean (Glycine max (L.) Merr.) cv. Forrest seeds were grown in greenhouse for fourteen days. Nuclei were isolated and embedded in agarose, restriction digested with Hind III BamHI or EcoRI, large size DNA fragments were ligated in vector V41 (pCLD04541) and electro transformed in DH10a cells. About 90,000 clones from BAC libraries were fingerprinted with HindIII and Hae III. Version 2 (automatic build) Contigs were built from 78,001 fingerprints. Contigs were manually examined to find the best non redundant tile path through the contigs representing about 13,000 clones. The clones were end sequenced."

Query Match 4.3%; Score 92.6; DB 29; Length 977;

Best Local Similarity 54.3%; Pred. No. 1.9e-12;

Matches 208; Conservative 0; Mismatches 174; Indels 1; Gaps 1;

QY 1 GGATCCCAACTTTTAGGAATGATCTTAAATTTTAGTTATAGTTCAAGTTAGAAAA 60

Db 572 GGTGTCCATCTCTACTTAAACATCTTAATCTTTATTATGATATCTTCATTGGAAA 513

QY 61 TCTTTACCAAGAGCTTTGAGTCCATTTGATGACATCCGTGAAACGGTGTA-CATGTCCTCG 119

Db 512 ATTTTCCCTTAAAGATGCAAGATGATTTATATATGTTGTAACCTCTTTGCATGCTTGT 453

QY 120 ATGGACTCACTTGGTTTCATTCGAAAAGTTCGAAAGTGTCATGAATATTTGTTG 179

Db 452 ATACCTTCATTTGAATTCATCTTAATAATTCATCTTATGAGTTAAGTATTTATCTTA 393

examined to find the best non redundant tile path through the contigs representing about 13,000 clones. The clones were end sequenced."

ORIGIN

Query Match	4.3%;	Score 92;	DB 29;	Length 832;	
Best Local Similarity	56.1%;	Pred. No. 2.7e-12;			
Matches 193;	Conservative 0;	Mismatches 150;	Indels 1;	Gaps 1;	
QY	19	ATGCACTCTAAATTTTGGTTATTAAGTTTCAAAGTTAGAAAATCTTTACCAAGAGCTTTG	78		
Db	171				
QY	79	AGTCCATTGATGACATCCGTGAAAC-GGTGTACATGTCCTCGATGGACTCACTTGGTTTC	137		
Db	231				
QY	138	ATTCGGAAAAGTTCGAAAGATGCATAAGAATATTGATTTTGGATCTTTCACTCGGTTG	197		
Db	291				
QY	198	GTGCGTTTCATCGAGTGACCTCAAGAGTCCTCAAAATATCAAAGCGGAATCAAAATTGAA	257		
Db	351				
QY	258	ATGTGATTGAATTCATTTTCTCTAAATGCACAAAACAGGGCATTCATAGCCTTTGTGTTT	317		
Db	411				
QY	471	TATTGAACTAAATCTTTTCTCTCTCTCACTCCATTAATCCATAGG	514		
Db	318	AAAGCAAAAACATTTCTTCCGATTCATCCCATTCGCTCATCGG	361		
QY	471	TATTGAACTAAATCTTTTCTCTCTCTCACTCCATTAATCCATAGG	514		
Db	411	ACCCTAAGATATTCATCTATTCCCTAATGTAGATGAATATTTTGGCTTTTAAATTG	470		

RESULT_5				
CG817604/c				GSS 18-NOV-2003
LOCUS	CG817604	821 bp	DNA	linear
DEFINITION	SOYDQ03.TV LargeInsertSoybeanGlib Glycine max genomic clone H3A21:MTP13A5, genomic survey sequence.			

CG817604	CG817604.1	GI:38272539
ACCESSION	VERSION	
KEYWORDS	GSS.	
SOURCE	Glycine max (soybean)	
ORGANISM	Glycine max	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.	

REFERENCE
1 (bases 1 to 821)
AUTHORS Shultz, J., Møksen, K., Shetty, J., Town, C.D., Koo, H., Potter, J., Wakefield, K., Zhang, H., Wu, C. and Lightfoot, D.A.
TITLE End sequencing of BACs comprising a provisional minimal tiling path from a fingerprint physical map of soybean (Glycine max) cultivar Forrest

JOURNAL
 COMMENT
 Unpublished (2003)
 Other_GSSs: SOYD003TH
 Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
 The Center of Excellence in Soybean Research, Teaching and
 Outreach, Southern Illinois University at Carbondale and Plant
 Genomics, The Institute for Genomic Research
 Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
 USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 618 453 1797
 Email: ga4082@siu.edu, jlschultz@siu.edu, cdtown@tigr.org (URL:
 http://bioinformatics.siu.edu)
 Clones approximating a minimum tiling path were re-arrayed from the
 library master plates prior to sequencing.
 For purposes of clone identification each clone name is a
 concatenation of the original clone location and its new location
 in the re-arrayed sequencing plates.
 Seq primer: GTAATACGACTCACTATAGGCG
 Class: BAC ends.

FEATURES

source

```

1. 821
/organism="Glycine max"
/moi_type="genomic DNA"
/cultivar="Forrest"
/db_xref="taxon:3847"
/clone="H32A21:MTp13A5"
/clone_lib="LargestInsertSoybeanGenLib"
/note="organ: leaves; Vector: pCLD04541 (pBELOBAC11 EcoRI
clones); Site 1: BstXI; Soybean (Glycine max (L.) Merr.)
cv. Forrest seeds were grown in greenhouse for fourteen
days. Nuclei were isolated and embedded in agarose,
restriction digested with Hind III BamHI or EcoRI, large
size DNA fragments were ligated in vector V41 (pCLD04541)
and electro transformed in DH10a cells. About 90,000
clones from BAC libraries were fingerprinted with HindIII
and Hae III. Version 2 (automatic build) contigs were
built from 78,001 fingerprints. Contigs were manually
examined to find the best non redundant tile path through
the contigs representing about 13,000 clones. The clones
were end sequenced."

```

ORIGIN

Query Match	4.2%;	Score	91.4;	DB	29;	Length	821;
Best Local Similarity	55.5%;	Pred.	No. 3.9e-12;				
Matches	196;	Conservative	0;	Mismatches	156;	Indels	1;
Gaps	1;						

Qy	1	GGATCCCAACTTTT	GAGGATCGATCTTAA	ATTTT	TAGTATTAAGTTC	CAAAAGTTG	AGAAAA	60
Db	571	GGTTGCCATCTCT	TAATTAAGACATCTT	PAGACIT	TATTTACTAGATCTT	CATTTT	GAAAA	512
Qy	61	TCCTTACCAAGAGC	TTTGAGTCCATTGATGACAT	CCGTGAAACGGT	GTGA-CATGTC	TCGCG	119	
Db	511	GTTTTTCCTAGAG	ATGCAAGATGATTAAC	TATGTGTGAACCTCTTT	TGCA	TGCTGTG	452	
Qy	120	ATGGACTCACTGTG	TTTCATTCGAAAGTT	CGAAAGTGCAT	TAAGAAAT	TGATTTG	179	
Db	451	ATACTTTTATT	TACATTCATCCT	AAAAAGTTCA	TATTCAGAGT	TAAAGTGT	392	
Qy	180	GATTCCTTCACTCG	GTGGTGGCTTCATGAGT	GACCTCAAGAGT	CTCCAAAT	TCAAAA	239	
Db	391	GATCTTTTAA	CATCTGTGTGGCTTC	CAATGTTACT	TGTATGCC	CATATCC	332	
Qy	240	GCCGAATCA	CAAAATGGAATGTG	ATGAAATTCATTTT	TGCTAAATGC	AAAAACAGG	299	
Db	331	GCACITTTA	CATTTTGAACCC	TAAAGTATTCATCT	ATTCCTAAGG	CAGATG	272	
Qy	300	TTCA	TAGCCTTTGTTT	TAAAGCAAAA	CACTCTTCTCCG	ATTCATCC	352	
Db	271	TTTTTGGC	CTTTTAAATATAT	TGTGACTAATCT	CTTTTCTCCTCAGT	CCAAATC	219	

RESULT 6

CG822306/c	CG822306	776 bp	DNA	linear	GSS 18-NOV-2001
LOCUS	SOYE21TV	LargeInsertSoybeanLib	Glycine max	genomic clone	
DEFINITION	H45P23:WTP15C17, genomic survey sequence.				

CG822306	CG822306.1	GI:38281855
ACCESSION	VERSION	
KEYWORDS	GSS	
SOURCE	Glycine max (soybean)	
ORGANISM	Glycine max	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euphorbia; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.	

REFERENCE
1. J. (bases 1 to 776)
AUTHORS
Shultz, J., Meksem, K., Shetty, J., Town, C.D., Koo, H., Potter, J.,
Wakefield, K., Zhang, H., Wu, C. and Lighfoot, D.A.
TITLE
End sequencing of BACs comprising a provisional minimal tiling pat-
tern from a fingerprint physical map of soybean (Glycine max) cultivar
Forrest
JOURNAL
Unpublished (2003)
COMMENT
Other GSSs: SOYEA21TH

Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
The Center of Excellence in Soybean Research, Teaching and
Outreach, Southern Illinois University at Carbondale and Plant
Genomics, The Institute for Genomic Research
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 618 453 1797
Email: g4082@siu.edu, jshultz@siu.edu, cdtown@tigr.org (URL:
http://bioinformatics.siu.edu)

Clones approximating a minimum tiling path were re-arrayed from the
library master plates prior to sequencing.
For purposes of clone identification each clone name is a
concatenation of the original clone location and its new location
in the re-arrayed sequencing plates.
Seq primer: GTAATACGACTCACTATAGGCG
Class: BAC ends

FEATURES

source
1. .776
/organism="Glycine max"
/mol_type="genomic DNA"
/cultivar="Forrest"
/db_xref="taxon:3847"
/clone="H45P23:WTP15C17"
/clone_lib="LargeInsertSoybeanGenLib"

/notes="Organ: Leaves; Vector: pCLD04541 (pBELOBAC11 EcorI
clones); Site 1: BstXI; Soybean (Glycine max (L.) Merr.)
cv. Forrest seeds were grown in greenhouse for fourteen
days. Nuclei were isolated and embedded in agarose,
restriction digested with Hind III BamHI or EcoRI, large
size DNA fragments were ligated in vector V41 (pCLD04541)
and electro transformed in DH10a cells. About 90,000
clones from BAC libraries were fingerprinted with HindIII
and Hae III. Version 2 (automatic build) Contigs were
built from 78,001 fingerprints. Contigs were manually
examined to find the best non redundant tiling path through
the contigs representing about 13,000 clones. The clones
were end sequenced."

ORIGIN

Query Match 4.2%; Score 89.8; DB 29; Length 776;

Best Local Similarity 55.2%; Pred. No. 9.9e-12;

Matches 195; Conservative 0; Mismatches 157; Indels 1; Gaps 1;

QY 1 GGATCCCACTTTTAGGAATGGATCTTAAATTTTATAGTTTCAAGTTAGAAAA 60

Db 568 GGTTCGCAATCTCTACTTAACATCTTAACATTTATTTATGAGATCCTGTTTGA 509

QY 61 TCTTTACCAAGACCTTTGAGTCCATTCATGACATCCGCGAAAC-GGTGTACATGTC 119

Db 508 GTTCTTCTTAATGATCAAGATGATTAACTATATGTGTGAATCTTTCTGATATCT 449

QY 120 ATGGACTCACTTGGTTTCATTCGGAAAAAGTTCGAAAGATGCATAGAAATTTG 179

Db 448 ATACTTTCAATTTGCTTCATCTTAATAATATCATCTAGTTAAAGTATTATCT 389

QY 180 GATCTTTCACTCGTGGTGGCTTCATGAGTGACCTCAAGAGTCTCCAAATATCA 239

Db 388 GATCTTTGACATCACTGTTTCCTCATGTGTAATCTTTAGATGTCCTCATATCT 329

QY 240 GCGGAATCAAAATTGAATGTGAATTCATTTTGTCTAATGCCAACAACAGGGCA 299

Db 328 GCACCTTTCAATTTGAAACCCCTTAAATATTCATCCATTCCTAGGCGAGAAGTA 269

QY 300 TCCATAGCTTTGTTTAAAGCAAAAACATCTTCTCGATTCATCCCATTC 352

Db 268 TTTTGTAGCTTTAAATATATCTGCATATCTCTTTTCATCTTCAGACCATC 216

RESULT 7

BH794864/c

LOCUS

DEFINITION BH794864 732 bp DNA linear GSS 02-APR-2002
ME_MBa0003J13f Manihot esculenta Manihot esculenta genomic clone
ME_MBa0003J13f, genomic survey sequence.

ACCESSION

VERSION BH794864.1 GI:19893414

KEYWORDS

SOURCE GSS.

ORGANISM

Manihot esculenta (cassava)
Manihot esculenta
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae;
Manihoteae; Manihot.

REFERENCE

AUTHORS

Tomkins,J.P., Fregene,M., Main,D., Goicoechea,J.L., Blackmon,B.,

TITLE

New Genomic Resources for Cassava (Manihot esculenta): Development
of a Deep-Coverage BAC Library and Preliminary STC Analysis

JOURNAL

COMMENT

Unpublished (2002)
Contact: Tomkins J

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson University, Clemson, SC 29634, USA

Tel: 864 656 6419

Fax: 864 656 4293

Email: jtmkns@clemson.edu

Total High Quality bases = 283

Seq primer: TAATACGACTCACTATAGGG

Class: BAC ends

High quality sequence start: 67

High quality sequence stop: 714.

FEATURES

source

1. .732

/organism="Manihot esculenta"

/mol_type="genomic DNA"

/strain="MECW72"

/db_xref="taxon:3983"

/clone="ME_MBa0003J13f"

/tissue_type="Leaf"

/lab_host="E. coli"

/clone_lib="Manihot esculenta"

/note="Vector: pCUGIBAC-1; Site 1: HindIII; Site 2: NotI;

For more details on library preparation and sequence

analysis see

http://www.genome.clemson.edu/projects/stc/cassava/ME_MBa

To order clones from this library see

http://www.genome.clemson.edu/orders "

ORIGIN

Query Match

Best Local Similarity 3.8%; Score 83; DB 28; Length 732;

Matches 189; Conservative 0; Mismatches 140; Indels 2; Gaps 2;

QY

4 TCCCACTTTTAGGAATGGATCTTAAATTTTATAGTTTCAAGTTAGAAAAATCT 63

Db 452 TCTCAAGATTTTGACAGTGTATCTTAAATTTTCTTCACTAGCGTAGCTTCTTCAAAGTTT 393

QY 64 TTACCAAGAGCTTTCAGTCCATTTGATGCATCCGTGAA-ACGGTGTACATGTCCTCGATG 122

Db 392 TTCCTAGTAGCTTTTAGAGATTTAGTAGATTAAATATATAGTACTATTTCACACA 333

QY 123 GACTCATTGGTTTCAT-TCGGAAGAGTTTCAAGAGTGCATAAGAATATTGATTTTGA 181

Db 332 GTTTCGCTGATTTTCATCTCAATTAATTTAGAAATTCATGACCCAGCAAGTTGCGTTTGA 273

QY 182 TTCCTTTCATCGTTGGTGGCTTCATGAGTGCCTCAAGAGTCCCTCCAAATATCAAAAGC 241

Db 272 TTCCTTTCGATTTGGGGTATTTCATAAGTCACTTTGGAGTTTATCCATATTTCTTTGGT 213

QY 242 CGAATCACAAATTGAAATGCTGATTCGAATTCATTTTGTCTATGTCACACAAACAGGCATT 301

Db 212 GGACTCCCACTTAAACAGGATTATATTCAGTTCATCCAAAGCCCAATTAAGTATTT 153

QY 302 CATAGCCTTTGTTTAAAGCAAAAACATTC 332

Db 152 TAAACCTTTAGCATTGGTGGAAATATTTTC 122

```

RESULT 8
CG824360/c
LOCUS
DEFINITION
    CG824360          791 bp      DNA          linear      GSS 18-NOV-2003
    SOYD062TH LargeInsertSoybeanGenLib Glycine max genomic clone
    H31P06:WTP13K3, genomic survey sequence.
ACCESSION
CG824360
VERSION
CG824360.1 GI:38298587
KEYWORDS
GSS.
SOURCE
    Glycine max (soybean)
ORGANISM
    Glycine max
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
    Glycine.
REFERENCE
    1 (bases 1 to 791)
    Shultz,J., Meksem,K., Shetty,J., Town,C.D., Koo,H., Potter,J.,
    Wakefield,K., Zhang,H., Wu,C. and Lightfoot,D.A.
    End sequencing of BACs comprising a provisional minimal tiling path
    from a fingerprint physical map of soybean (Glycine max) cultivar
    Forrest
JOURNAL
    Unpublished (2003)
COMMENT
    Other GSSs: SOYDQ62TV
    Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
    The Center of Excellence in Soybean Research, Teaching and
    Outreach, Southern Illinois University at Carbondale and Plant
    Genomics, The Institute for Genomic Research
    Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
    USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 618 453 1797
    Email: ga4082@siu.edu, jlshultz@siu.edu, cdtown@igr.org (URL:
    http://bioinformatics.siu.edu)
    Clones approximating a minimum tiling path were re-arrayed from the
    library master plates prior to sequencing.
    For purposes of clone identification each clone name is a
    concatenation of the original clone location and its new location
    in the re-arrayed sequencing plates.
    Seq primer: AATTAACCTCACTAAGGG
    Class: BAC ends.

FEATURES
    Location/Qualifiers
    1..791
       /organism="Glycine max"
       /mol_type="genomic DNA"
       /cultivar="Forrest"
       /db_xref="taxon:3847"
       /clone="H31P06:WTP13K3"
       /clone_lib="LargeInsertSoybeanGenLib"
       /notes="Organ: Leaves; Vector: pCLD04541 (pBELOBAC11 EcorI
       clones); Site 1: BstXI; Soybean (Glycine max (L.) Merr.)
       cv. Forrest seeds were grown in greenhouse for fourteen
       days. Nuclei were isolated and embedded in agarose, large
       restriction digested with Hind III BamHI or EcoRI, large
       size DNA fragments were ligated in vector V41 (pCLD04541)
       and electro transformed in DH10a cells. About 90,000
       clones from BAC libraries were fingerprinted with HindIII
       and Hae III. Version 2 (automatic build) Contigs were
       built from 78,001 fingerprints. Contigs were manually
       examined to find the best non redundant tile path through
       the contigs representing about 13,000 clones. The clones
       were end sequenced."

ORIGIN
    Query Match          3.8%; Score 83; DB 29; Length 791;
    Best Local Similarity 54.1%; Pred. No. 5.3e-10;
    Matches 190; Conservative 0; Mismatches 160; Indels 1; Gaps 1;

    QY      1  GGATCCCAACTTTTAGGAATGCATCTTAAATTTTAGTTATAGTTCAAGTTAGAAAA 60
    Db      603  GGTGCCCATCTCTACTTAAACATCTTACACCTGTTTATAGATCTTCATTTGAAT 544
    QY      61  TCTTTACCAAGACGTTTGAGTCCATGTGATGACATCCGGAACCGGTGACATGTCGCG- 119
    Db      543  TCTTTGCTAGGCTGCTAGATGATTACTATATGTGTAATCTCTTTTGCATACCTCTGA 484

```

```

QY      120  ATGGACTCACTTGGTTTCATTCGAAAAAGTTGAAAGAGTGCAAGAAATATTGATTG 179
Db      483  ATATTTTCATTTGCATTCATCTTAATAATTAATCATCTAGCTAGTAGTCATTTACCTA 424
QY      180  GATCTTTTCACATCGGTTGGTTCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAA 239
Db      423  GATCTTTTAACATCTGTAGTTCTTTCATGTTTAATCGAAGAGTGTCCACATTTCTCTTA 364
QY      240  GCCGAATCAAAATTTGAATGTGATTGAATTCATTTTGTCTTAATGACAAACAGGGCA 299
Db      363  GCATCTTCTACATTTTGAACCCCTGAAGTATTCATTCCTTAGGGCAGATGTTATTATG 304
QY      300  TTCATAGCCTTTGTGTTTAAAGCAAAACATTTCTTCGCAATCATCCCAT 350
Db      303  TTTTGGCTTTTAGTTGTTGTTACTCGCTTTTCTATCCTCTTCAGACCAT 253

RESULT 9
CG817563
LOCUS
DEFINITION
    CG817563          890 bp      DNA          linear      GSS 18-NOV-2003
    SOYBC57TV LargeInsertSoybeanGenLib Glycine max genomic clone
    B50J09:MTP7J17, genomic survey sequence.
ACCESSION
CG817563
VERSION
CG817563.1 GI:38272458
KEYWORDS
GSS.
SOURCE
    Glycine max (soybean)
ORGANISM
    Glycine max
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
    Glycine.
REFERENCE
    1 (bases 1 to 890)
    Shultz,J., Meksem,K., Shetty,J., Town,C.D., Koo,H., Potter,J.,
    Wakefield,K., Zhang,H., Wu,C. and Lightfoot,D.A.
    End sequencing of BACs comprising a provisional minimal tiling path
    from a fingerprint physical map of soybean (Glycine max) cultivar
    Forrest
JOURNAL
    Unpublished (2003)
COMMENT
    Other GSSs: SOYBC57TH
    Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
    The Center of Excellence in Soybean Research, Teaching and
    Outreach, Southern Illinois University at Carbondale and Plant
    Genomics, The Institute for Genomic Research
    Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
    USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
    Tel: 618 453 1797
    Email: ga4082@siu.edu, jlshultz@siu.edu, cdtown@igr.org (URL:
    http://bioinformatics.siu.edu)
    Clones approximating a minimum tiling path were re-arrayed from the
    library master plates prior to sequencing.
    For purposes of clone identification each clone name is a
    concatenation of the original clone location and its new location
    in the re-arrayed sequencing plates.
    Seq primer: GTAATACGACTCACTATAGGGC
    Class: BAC ends.

FEATURES
    Location/Qualifiers
    1..890
       /organism="Glycine max"
       /mol_type="genomic DNA"
       /cultivar="Forrest"
       /db_xref="taxon:3847"
       /clone="B50J09:MTP7J17"
       /clone_lib="LargeInsertSoybeanGenLib"
       /notes="Organ: Leaves; Vector: pCLD04541 (pBELOBAC11 EcorI
       clones); Site 1: BstXI; Soybean (Glycine max (L.) Merr.)
       cv. Forrest seeds were grown in greenhouse for fourteen
       days. Nuclei were isolated and embedded in agarose,
       restriction digested with Hind III BamHI or EcoRI, large
       size DNA fragments were ligated in vector V41 (pCLD04541)
       and electro transformed in DH10a cells. About 90,000
       clones from BAC libraries were fingerprinted with HindIII
       and Hae III. Version 2 (automatic build) Contigs were

```


built from 78,001 fingerprints. Contigs were manually examined to find the best non redundant tile path through the contigs representing about 13,000 clones. The clones were end sequenced."

ORIGIN

```

Query Match      3.8%; Score 81.8; DB 29; Length 890;
Best Local Similarity 53.8%; Pred. No. 1.1e-09;
Matches 190; Conservative 0; Mismatches 162; Indels 1; Gaps 1;

QY 1 GGATCCCACTTTAGGAATGGATCTTAAATTTAGTTATAGTTCAAGATTTAGAAAAA 60
Db 412 GGTTCCTCAAGTTCTTGAAGATAGTTGAGGATTTGATGTTTAGCTCATCGGTTCAAA 471
QY 61 TCTTTACCAAGAGTTTGATGTCATTCATGATGATCGGTGAAC-GGTGTACATGTCCTCG 119
Db 472 ATTTTCCAGACCAAGAAGGTGATTCACATATGGGTGAATCTGTGGAAGTTCTAAA 531
QY 120 ATGAGTCACTTGGTTTCATTCGGAAGAAGTTTGAAGAAGTGCATAAGAAATATTGATTTG 179
Db 532 ATGTTTCTCCATTTCTTCAATTCGAAAGCTTCATATTGGATAGATAGTGTGTTCTT 591
QY 180 GATCTTTCATCGGTGGTGGCTTCATGAGTGACCTCAAGAGTCTCTCAAAATATCAAAA 239
Db 592 GATCTTCTTACATCTCTAGAGCCCTTCATGAGTCACTTCAAGCATTTTCCATATTTCTTT 651
QY 240 GCCGAATCAAAATTTGAATGTGATTTGATTTGTTGCTAATGCACAAAACAGGGCA 299
Db 652 GCATTTTGTACCTTCAGTATGGAAGAATTCATCAGAGATTAACCAAGATTTAAATG 711
QY 300 TCCATAGCCTTTGTTTAAAGCAAAACATTTCTTCGATTCATCCCATTC 352
Db 712 TTTTAGCCTTTTGATCATCTTTAACTATTCTTTCTCATCATCTTTCATC 764

```

RESULT 10

```

CG820967/c
LOCUS
DEFINITION
  SOYCD687V LargeInsertSoybeanGenLib Glycine max genomic clone
  H05C12-MTP8L16, genomic survey sequence.
ACCESSION
  CG820967
VERSION
  CG820967.1 GI:38279201
KEYWORDS
  GSS.
SOURCE
  Glycine max (soybean)
  Glycine max
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
REFERENCE
  1 (bases 1 to 927)
  Shultz,J., Meksem,K., Shetty,J., Town,C.D., Koo,H., Potter,J.,
  Wakefield,K., Zhang,H., Wu,C. and Lightfoot,D.A.
  End sequencing of BACs comprising a provisional minimal tiling path
  from a fingerprint physical map of soybean (Glycine max) cultivar
  Forrest
  Unpublished (2003)
  Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
  The Center of Excellence in Soybean Research, Teaching and
  Outreach, Southern Illinois University at Carbondale and Plant
  Genomics, The Institute for Genomic Research
  Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
  USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 618 453 1797
  Email: ga4082@siu.edu, jshultz@siu.edu, cdtown@tigr.org (URL:
  http://bioinformatics.siu.edu)
  Clones approximating a minimum tiling path were re-arrayed from the
  library master plates prior to sequencing.
  For purposes of clone identification each clone name is a
  concatenation of the original clone location and its new location
  in the re-arrayed sequencing plates.
  Seq primer: GTAATACGACTCACTATAGGGC
  Class: BAC ends.
  Location/Qualifiers

```

FEATURES

source

```

1. 927
/organism="Glycine max"
/mol_type="genomic DNA"
/cultivar="Forrest"
/db_xref="taxon:3847"
/clone="H05C12-MTP8L16"
/clone_lib="LargeInsertSoybeanGenLib"
/notes="Organ: Leaves; Vector: pCLD04541 (pBELOBAC11 EcorI
clones); Site 1: BatXi; Soybean (Glycine max (L.) Merr.)
cv. Forrest seeds were grown in greenhouse for fourteen
days. Nuclei were isolated and embedded in agarose,
restriction digested with Hind III BamHI or EcoRI, large
size DNA fragments were ligated in vector V41 (pCLD04541)
and electro transformed in DH10a cells. About 90,000
clones from BAC libraries were fingerprinted with HindIII
and Hae III. Version 2 (automatic build) Contigs were
built from 78,001 fingerprints. Contigs were manually
examined to find the best non redundant tile path through
the contigs representing about 13,000 clones. The clones
were end sequenced."

```

ORIGIN

```

Query Match      3.8%; Score 81.8; DB 29; Length 927;
Best Local Similarity 55.0%; Pred. No. 1.1e-09;
Matches 181; Conservative 0; Mismatches 147; Indels 1; Gaps 1;

QY 23 ATCTTAAATTTTAGTTTAAAGTTCAAGTTTCAAGAAATCTTTTACCAAGAGCTTTGAGTC 82
Db 579 ATCTTAAATCTTTGTTTATGAGATCTTCATTTTGAAATCTTTTCCCTAAGGCTGCAAGAT 520
QY 83 CATTGATGATCCGCTGGAAC-GGTGTACATCTCTCCGATGAGTCACTTGGTTTCATTC 141
Db 519 GATTTACTATATGTGTAAACCTCTTTTGCAATGCTTGAATGCTTTTCATTTGCAATTCAC 460
QY 142 GGAAGAAGTTTGAAGAGTGCATAGAATATTGATTTGGATCTTTTCATCGGTGGTGC 201
Db 459 TAAATAGTTTCATATTCATGAGTTAATGTTTATCTAGATCTTTTAAATCTATTTGTC 400
QY 202 CTTTATGAGTACCTCAAGAGTCTCTCAAAATATCAAAAGCCGAATCACAATTTGAAATGT 261
Db 399 CTTTATGTTTACTTGTAGAGTGTCCACATCTCTTAGCACTCTTTCACATTTGAACCC 340
QY 262 GATTGAATTCATTTTGTCTATGACACAAAACAGGGCATTCATAGCTTTGTTTAAAG 321
Db 339 TGAATATTCATCCATTCCTAGGAGATGTAATATGTTTGGCTTTTAAATATATT 280
QY 322 CAAAACATTTCTTCTCCGATTCATCCCAT 350
Db 279 TTAATCGTCTTCTATACTCTTCAGACCAT 251

```

RESULT 11

```

CG815315/c
LOCUS
DEFINITION
  SOYED29TV LargeInsertSoybeanGenLib Glycine max genomic clone
  H50C23-MTP15F10, genomic survey sequence.
ACCESSION
  CG815315
VERSION
  CG815315.1 GI:38268789
KEYWORDS
  GSS.
SOURCE
  Glycine max (soybean)
  Glycine max
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
  Glycine.
REFERENCE
  1 (bases 1 to 733)
  Shultz,J., Meksem,K., Shetty,J., Town,C.D., Koo,H., Potter,J.,
  Wakefield,K., Zhang,H., Wu,C. and Lightfoot,D.A.
  End sequencing of BACs comprising a provisional minimal tiling path
  from a fingerprint physical map of soybean (Glycine max) cultivar
  Forrest
  Unpublished (2003)
  Other_GSSs: SOYED29TH

```

Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
The Center of Excellence in Soybean Research, Teaching and
Outreach, Southern Illinois University at Carbondale and Plant
Genomics, The Institute for Genomic Research
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 618 453 1797
Email: ga4082@siu.edu, jshultz@siu.edu, cdtown@tigr.org (URL:
http://bioinformatics.siu.edu)
Clones approximating a minimum tiling path were re-arrayed from the
library master plates prior to sequencing.
For purposes of clone identification each clone name is a
concatenation of the original clone location and its new location
in the re-arrayed sequencing plates.
Seq primer: GTAATACGACTCACTATAGGCG
Class: BAC ends.

FEATURES
source
1..733
/organism="Glycine max"
/mol_type="genomic DNA"
/cultivar="Forrest"
/db_xref="taxon:3847"
/clone="H50C23:WTP15F10"
/clone_lib="LargeInsertSoybeanGenLib"
/notes="Organ: Leaves; Vector: pCLD04541 (pBEOBAC11 BcoRI
clones); Site 1: BstXI; Soybean (Glycine max (L.) Merr.)
cv. Forrest seeds were grown in greenhouse for fourteen
days. Nuclei were isolated and embedded in agarose,
restriction digested with Hind III BamHI or EcoRI, large
size DNA fragments were ligated in vector v41 (pCLD04541)
and electro transformed in DH10a cells. About 90,000
clones from BAC libraries were fingerprinted with HindIII
and Hae III. Version 2 (automatic build) Contigs were
built from 78,001 fingerprints. Contigs were manually
examined to find the best non redundant tile path through
the contigs representing about 13,000 clones. The clones
were end sequenced."

ORIGIN

Query Match 3.8%; Score 81.4; DB 29; Length 733;
Best Local Similarity 53.8%; Pred. No. 1.4e-09;
Matches 189; Conservative 0; Mismatches 161; Indels 1; Gaps 1;
Qy 1 GGATCCCAACTTTAGGAAGGATCTTAAATTTTAGTTATTAAGTTCAAGTTAGAAAA 60
Db 603 GGTGGCCATCTCTACTTAAACATCTTAAACCTTGTTTTAAAGATCTCTCAATTTGAAAT 544
Qy 61 TCTTTACCAAGAGCTTTGAGTCCATTTGATGACATCCGTTGAAC-GGTGTACATGTCTCCG 119
Db 543 TCTTTGCTTAAGGCTGCTAGATGATTTACTATATGTGTAAATCTCTTTTGATCTTTGA 484
Qy 120 ATGGACTCAGTGTGTTTCATTCGGAAGAGTTGCGAAGAGTGATGATGAAGATTTGATTTG 179
Db 483 ATATTTTCATTTGCAATTCATTTAAATTAATTCATCTCATGATGATGTCATTTATCCTA 424
Qy 180 GATTTCTTCACTCGTTGGTGGCTTCAGAGTCCCTCAAGAGTCTCCBAAATCAAAA 239
Db 423 GATCTTTAAACATCTGTAGTTCCCTCATGTGTTAATTAAGAGTGTCACATTTTCCTTA 364
Qy 240 GCGGAATCACAAATTTGAATGTGATTTGAATTCATTTTGTCTATGCAACAAACAGGCA 299
Db 363 GCATCTTACAAATTTGAACCCCTGAAATATTCATCCATTCCTAGGCGAGATGTTATTATG 304
Qy 300 TCCATAGCCTTTGTTTAAAGCAAAAACATCTCTCCGATTCATCCAT 350
Db 303 TTTTGGCTTTTAAAGTTGTTATTTGACTCATTTTCTATCTTCTTCTCAGACCAT 253

RESULT 12
CG817875/c
LOCUS
DEFINITION
H11E21:WTP9H13, genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CG817875
CG817875.1 GI:38273079
GSS.
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 872)
Shultz,J., Mekeem,K., Shetty,J., Town,C.D., Koo,H., Potter,J.,
Wakefield,K., Zhang,H., Wu,C. and Lightfoot,D.A.
End sequencing of BACs comprising a provisional minimal tiling path
from a fingerprint physical map of soybean (Glycine max) cultivar
Forrest
Unpublished (2003)
Other GSSs: SOYDC43TH
Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
The Center of Excellence in Soybean Research, Teaching and
Outreach, Southern Illinois University at Carbondale and Plant
Genomics, The Institute for Genomic Research
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 618 453 1797
Email: ga4082@siu.edu, jshultz@siu.edu, cdtown@tigr.org (URL:
http://bioinformatics.siu.edu)
Clones approximating a minimum tiling path were re-arrayed from the
library master plates prior to sequencing.
For purposes of clone identification each clone name is a
concatenation of the original clone location and its new location
in the re-arrayed sequencing plates.
Seq primer: GTAATACGACTCACTATAGGCG
Class: BAC ends.

Location/Qualifiers
1..872
/organism="Glycine max"
/mol_type="genomic DNA"
/cultivar="Forrest"
/db_xref="taxon:3847"
/clone="H11E21:WTP9H13"
/clone_lib="LargeInsertSoybeanGenLib"
/note="Organ: Leaves; Vector: pCLD04541 (pBEOBAC11 BcoRI
clones); Site 1: BstXI; Soybean (Glycine max (L.) Merr.)
cv. Forrest seeds were grown in greenhouse for fourteen
days. Nuclei were isolated and embedded in agarose,
restriction digested with Hind III BamHI or EcoRI, large
size DNA fragments were ligated in vector v41 (pCLD04541)
and electro transformed in DH10a cells. About 90,000
clones from BAC libraries were fingerprinted with HindIII
and Hae III. Version 2 (automatic build) Contigs were
built from 78,001 fingerprints. Contigs were manually
examined to find the best non redundant tile path through
the contigs representing about 13,000 clones. The clones
were end sequenced."

ORIGIN

Query Match 3.8%; Score 81.4; DB 29; Length 872;
Best Local Similarity 53.8%; Pred. No. 1.4e-09;
Matches 189; Conservative 0; Mismatches 161; Indels 1; Gaps 1;
Qy 1 GGATCCCAACTTTTAGGAATGGATCTTAAATTTTAGTTATTAAGTTCAAGTTAGAAAA 60
Db 596 GGTGGCCATCTCTACTTAAACATCTTAAACCTTGTTTTAAAGATCTCTCAATTTGAAAT 537
Qy 61 TCTTTACCAAGAGCTTTGAGTCCATTTGATGACATCCGTTGAACGGTGTACATGTCTCCG- 119
Db 536 TCTTTGCGCTAAGGCTGCTAGATGATTTACTATATGTGTAAATCTCTTTTGGATCTCTGA 477
Qy 120 ATGGACTCAGTGTGTTTTCATTCGGAAGAGTTTCGAAAGAGTGCATAGAATATTGATTTTG 179
Db 476 ATATTTTCATTTGCAATTCATTTCTAAATTAATTCATCTCATGATGATGATTTATCCTTA 417
Qy 180 GATTTCTTCACTCGGTGGTGGTGGCTTCCTTCATGATGATGATGATGATGATGATGATGAT 239

```

Db      416  GATCTTTTAACTCTCTAGTTCCTTCATGTTTAAATCGAAGAGTGTCACACATTTCTCTTA 357
QY      240  GCGGATCAAAATGGAATGTAATCAATTTTGTCTTAATGCAAAACAGGGCA 299
Db      356  GCACTCTTAAATTTGAAACCTTGAATAATTCATCCATTCCTAGGCGAGATGTTATTATG 297
QY      300  TTCTATAGCCCTTGTGTTTAAAGCAAAACATCTTCTCGGATTCATCCCAT 350
Db      296  TTTTGGCTTTAAGTTGATTGTACTCGTTTCTATCCCTTCCTCGGACCAT 246

RESULT 13
LOCUS   CB340076
DEFINITION CB340076 295 bp mRNA linear EST 14-MAR-2003
          CA23EI02IRC_C04 Cabernet Sauvignon Leaf - CA23EI Vitis vinifera
          cDNA clone CA23EI02IRC_C04 3', mRNA sequence.
ACCESSION CB340076
VERSION   CB340076.1 GI:28960151
KEYWORDS  Vitis vinifera
SOURCE    Vitis vinifera
ORGANISM  Vitis vinifera
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 295)
AUTHORS   Goes da Silva,F., Lim,H., Iandolino,A., Baek,J., Jones,K.,
          Walker,M.A. and Cook,D.R.
TITLE      Transcriptional responses of Vitis vinifera to infection by the
          bacterial pathogen Xylella fastidiosa
JOURNAL    Unpublished (2003)
COMMENT    Contact: Douglas Cook, PhD
          CAES Genome Facility
          UC Davis, Plant Pathology
          One Shields Ave, Davis, CA 95616, USA
          Tel: 530 754 6561
          Fax: 530 754 6617
          Email: drcook@ucdavis.edu
          Seq primer: GCCAAGCAATGGTCTAG.

FEATURES             source
    1..295
         /organism="Vitis vinifera"
         /mol_type="mRNA"
         /cultivar="Cabernet Sauvignon"
         /db_xref="taxon:29760"
         /clone="CA23EI02IRC_C04"
         /sex="Hermaphrodite"
         /dev_stage="Mid-season leaf material, collected July 25,
         2001."
         /lab_host="DH5alpha"
         /clone_lib="Cabernet Sauvignon Leaf - CA23EI"
         /note="Organ: Leaf; Vector: pDNR; Site_1: SfiI; Site_2:
         SfiI; CA23EI is a cDNA library of Cabernet Sauvignon
         leaves. The leaves were collected on July 25, 2001, in
         Napa Valley, California, and represent leaves in
         mid-season development. These leaves were verified to be
         infected with the bacterial pathogen, Xylella fastidiosa,
         based on a diagnostic assay using PCR and Xylella-specific
         primer pairs. The plants were asymptomatic at the time of
         collection, but later developed symptoms. cDNAs were made
         by oligo-dT priming and directionally cloned. 5' and 3'
         adaptors were used in cloning as follows:
         5'-AAGCACTGTATCAACGACGATGGCATACGCGCGG-3' and
         5'-ATTCTAGAGCGGAGCGCGGCACATG-AT(30)NN-3'. Library was
         constructed using the Clontech Creator SMART kit and
         size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
Query Match          3.8%; Score 81.2; DB 14; Length 295;
Best Local Similarity 63.8%; Pred. No. 1.4e-09;
Matches 139; Conservative 0; Mismatches 78; Indels 1; Gaps 1;

QY      64  TTACCAAGAGCTTTGAGTCCATTGATGACATCCGTGAAC-GGTGTACATGCTCCGATG 122

```

```

Db      76  TTCCCAAGCTTGAAGTTCATTCACAATCCCATGAACCTAAAAACATCTCTACAATA 135
QY      123  GACTCACTTGGTTTCATTCGAAAGCTTCGAAAGAGTGCATAAGATATTGATTTGGAT 182
Db      136  GATTCAAGATCCTTTCATAAAAAAATAATTCAATAATGCGCAAGGATGTTGATTTAGAC 195
QY      183  TCTTTCACTCGGTTGGTGGCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAAGCC 242
Db      196  TCTTTAACTTGGTTGGTTCCTCATGAGTAATTTCTAATAGTCTCCAAATCTCTTAGCC 255
QY      243  GAATCAAAATTCAAATGTGATTGAATTCATTTTGTGTC 280
Db      256  AATTTCATTTGAAACATGATTAATTCATTTCTATC 293

RESULT 14
LOCUS   CG816026/c
DEFINITION CG816026 795 bp DNA linear GSS 18-NOV-2003
          SOYFH38TV LargeInsertSoybeanGenLib Glycine max genomic clone
          H41104:SEQ3H4, genomic survey sequence.
ACCESSION CG816026
VERSION   CG816026.1 GI:38269500
KEYWORDS  Glycine max (soybean)
SOURCE    Glycine max (soybean)
ORGANISM  Glycine max
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
          Glycine.
REFERENCE 1 (bases 1 to 795)
AUTHORS   Shultz,J., Meksem,K., Shetty,J., Town,C.D., Koo,H., Potter,J.,
          Wakefield,K., Zhang,H., Wu,C. and Lightfoot,D.A.
TITLE      End sequencing of BACs comprising a provisional minimal tiling path
          from a fingerprint physical map of soybean (Glycine max) cultivar
          Forrest
JOURNAL    Unpublished (2003)
COMMENT    Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
          The Center of Excellence in Soybean Research, Teaching and
          Outreach, Southern Illinois University at Carbondale and Plant
          Genomics, The Institute for Genomic Research
          Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62301-4415,
          USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 618 453 1797
          Email: ga4082@siu.edu, jlshultz@siu.edu, cdtown@tigr.org (URL:
          http://bioinformatics.siu.edu)
          Clones approximating a minimum tiling path were re-arrayed from the
          library master plates prior to sequencing.
          For purposes of clone identification each clone name is a
          concatenation of the original clone location and its new location
          in the re-arrayed sequencing plates.
          Seq primer: GTAATACGACTCACTATAGGGC
          Class: BAC ends.

FEATURES             Location/Qualifiers
    1..795
         /organism="Glycine max"
         /mol_type="genomic DNA"
         /cultivar="Forrest"
         /db_xref="taxon:3847"
         /clone="H41104:SEQ3H4"
         /clone_lib="LargeInsertSoybeanGenLib"
         /note="Organ: Leaves; Vector: pCLD04541 (pBETOBAC11 EcorI
         clones); Site_1: BstXI; Soybean (Glycine max (L.) Merr.)
         cv. Forrest seeds were grown in greenhouse for fourteen
         days. Nuclei were isolated and embedded in agarose,
         restriction digested with Hind III BamHI or EcorI, large
         size DNA fragments were ligated in vector V41 (pCLD04541)
         and electro transformed in DH10a cells. About 90,000
         clones from BAC libraries were fingerprinted with HindIII
         and Hae III. Version 2 (automatic build) Contigs were
         built from 78,001 fingerprints. Contigs were manually
         examined to find the best non redundant tile path through
         the contigs representing about 13,000 clones. The clones

```

were end sequenced."

/clone_lib="LargeInsertSoybeanGenLib"
/note="Organ: Leaves; Vector: pCLD04541 (pBELOBAC11 EcoRI
clones); Site 1: BstXI; Soybean (Glycine max (L.) Merr.)
cv. Forrest seeds were grown in greenhouse for fourteen
days. Nuclei were isolated and embedded in agarose,
restriction digested with Hind III BamHI or EcoRI, large
size DNA fragments were ligated in vector V41 (pCLD04541)
and electro transformed in DH10a cells. About 90,000
clones from BAC libraries were fingerprinted with HindIII
and Hae III. Version 2 (automatic build) Contigs were
built from 78,001 fingerprints. Contigs were manually
examined to find the best non redundant tile path through
the contigs representing about 13,000 clones. The clones
were end sequenced."

ORIGIN

Query Match 3.8%; Score 81.2; DB 29; Length 914;
Best Local Similarity 53.3%; Pred. No. 1.5e-09;
Matches 193; Conservative 0; Mismatches 168; Indels 1; Gaps 1;
QY 1 GGATCCCACTTTTAGGAATGGATCTTAAATTTTATAGTTTCAAGTTAGAAAA 60
DB 571 GGTGCGAATCTCTACTAGACATTTCAAACTTTTATTTATAGATCCTCATTTGGAAT 512
QY 61 TCTTTACCAAGAGCTTTGAGTCCATTCGATGACATCCGTGAAAC-GGTGTACATGTCTCCG 119
DB 511 ATTTTCTTAAAGATGCAAGATGATTAACTATGTGTGGAACCTCTTTTGCATGTCTGT 452
QY 120 ATGAGCTCACTTGGTTTCATTCGGAAAGTTGGAAGAGTGCAATAGATATTGATTTTG 179
DB 451 ATACTTTTCATTTGCATTCATCTTAAATAGTTTATTTATGAGTTAATGTGTTATCCTA 392
QY 180 GATTTTTCATCTCGTTTGGTCCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAA 239
DB 391 GATCTTTTAACTCTGTGTGGCCTTCATGTTTACTTTGAGGTATCCCAATATCTCTT 332
QY 240 GCCGAATCAAAATGAAATGTGATTGAATTCATTTTGTCTAATGCACAAAAACAGGCCA 299
DB 331 TCATTTTACAGTTTGCATACCTTAAAGTATTCATCCATTTTAGGGCAGATGTAATTATG 272
QY 300 TCCATAGCCTTTGTGTTTAAAGCAAAACATTTCTTCCGATTCATCCCATTCGCTCATC 359
DB 271 TTTTGGCTTTTAAAGTTGTATTGTACTAGTCTTTCTTTCTTCTTCTTCTTCTTCTTCC 212
QY 360 GG 361
DB 211 GG 210

Search completed: September 19, 2004, 02:37:42
Job time : 4975.1 secs

ORIGIN

Query Match 3.8%; Score 81.2; DB 29; Length 795;
Best Local Similarity 57.2%; Pred. No. 1.5e-09;
Matches 166; Conservative 0; Mismatches 123; Indels 1; Gaps 1;
QY 1 GGATCCCACTTTTAGGAATGGATCTTAAATTTTATAGTTTCAAGTTAGAAAA 60
DB 603 GGTGCGAATCTCTACTAGACATCTTAATACCTTTTATAGATCTTCATTTGAAAA 544
QY 61 TCTTTACCAAGAGCTTTGAGTCCATTCGATGACATCCGTGAAACGGTGTA-CATGTCTCCG 119
DB 543 TCTTTCCCTAAGCTGCAAGATGATTACTATATGTGTGAACTCTTTTGCATGTCTTGA 484
QY 120 ATGAGCTCACTTGGTTTCATTCGGAAAGTTGGAAGAGTGCAATAGATATTGATTTTG 179
DB 483 ATGCTTTTCATTTGCATTCATCTTAAATAGTTTATATTCATGAGTTAATGTGTTATCCTA 424
QY 180 GATTTTTCATCTCGTTTGGTCCCTTCATGAGTGACCTCAAGAGTCTCCAAATATCAAAA 239
DB 423 GATCTTTTAACTCTGTGTGGCCTTCATGTTTACTTTGAGGTATCCCAATATCTCTT 364
QY 240 GCCGAATCAAAATGAAATGTGATTGAATTCATTTTGTCTAATGCACA 289
DB 363 GCATCTTCAATTTGAAACCTTGAAATATTCATTCATTCCTCCAGGGCAGA 314

RESULT 15
CG822083/c

LOCUS 914 bp DNA linear GSS 18-NOV-2003
DEFINITION SOYD15TV LargeInsertSoybeanGenLib Glycine max genomic clone
H37B06:MTPL13D6, genomic survey sequence.

ACCESSION CG822083
VERSION CG822083.1 GI:38281410

KEYWORDS GSS.

SOURCE Glycine max (soybean)

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE 1 (bases 1 to 914)

AUTHORS Shultz, J., Meksem, K., Shetty, J., Town, C.D., Koo, H., Potter, J.,
Wakefield, K., Zhang, H., Wu, C. and Lightfoot, D.A.
TITLE End sequencing of BACs comprising a provisional minimal tiling path
from a fingerprint physical map of soybean (Glycine max) cultivar
Forrest

JOURNAL Unpublished (2003)

COMMENT

Other GSSs: SOYD15TH
Contact: Chris Town, J. L. Shultz and D. A. Lightfoot
The Center of Excellence in Soybean Research, Teaching and
Outreach, Southern Illinois University at Carbondale and Plant
Genomics, The Institute for Genomic Research
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 618 453 1797

Email: g4092@siu.edu, jshultz@siu.edu, cdtown@igrr.org (URL:
http://bioinformatics.siu.edu)

Clones approximating a minimum tiling path were re-arrayed from the
library master plates prior to sequencing.

For purposes of clone identification each clone name is a
concatenation of the original clone location and its new location
in the re-arrayed sequencing plates.

Seq primer: GTAATACGACTCACTATAGGCG

Class: BAC ends.

FEATURES

source

1. .914
Location/Qualifiers
/organism="Glycine max"
/mol_type="genomic DNA"
/cultivar="Forrest"
/db_xref="taxon:3847"
/clone="H37B06:MTPL13D6"